OPTIMAL CONTROL AND PARAMETER DETERMINATION IN VOLTERRA TYPE INTEGRAL AND DELAY DIFFERENTIAL SYSTEMS

By

NALINI KUMAR PATEL

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

MARCH, 1981

OPTIMAL CONTROL AND PARAMETER DETERMINATION IN VOLTERRA TYPE INTEGRAL AND DELAY DIFFERENTIAL SYSTEMS

A Thesis Submitted
in Partial Fulfilment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

By
NALINI KUMAR PATEL

to the
DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

MARCH, 1981

To loving

NISHANT

CERTIFICATE

It is certified that this work entitled 'OPTIMAL CONTROL AND PARAMETER DETERMINATION IN VOLTERRA TYPE INTEGRAL AND DELAY DIFFERENTIAL SYSTEMS' by Nalini Kumar Patel has been carried out under our supervision and that this work has not been submitted elsewhere for the award of a degree.

S. S. Prabhu.
Dr. S.S. Prabhu
Professor
Department of Electrical Engineering
Indian Institute of Technology
Kanpur.

PCDas
Dr. P.C. Das
Professor
Department of Mathematics
Indian Institute of Technology
Kanpur.

ACKNOWLEDGEMENT

I am very much indebted to my supervisors Dr. P.C. Das and Dr. S.S. Prabhu for their advice and keen interest during the course of this work.

I am highly grateful to Dr. L.P. Singh for his advice and help rendered in various problems.

Thanks are due to all my friends especially, K.K. Goyal, Sachchidanand, Munna, Omprakash and Kalyan for their cooperat in all spheres of life at Kanpur.

My sincere thanks are to Mr. C.M. Abraham for the care ta in typing, Mr. B.B. Srivastava for preparation of the drawings and Mr. Rajeswari for cyclostyling of this thesis.

Finally, I would like to thank my sister Jayanti and wife Shaila for obvious reasons.

nalini

SYNOPSIS

NALINI KUMAR PATEL
Department of Electrical Engineering
Indian Institute of Technology
Kanpur - 208016
March, 1981

OPTIMAL CONTROL AND PARAMETER DETERMINATION IN VOLTERRA TYPE INTEGRAL AND DELAY DIFFERENTIAL SYSTEMS

A wide variety of systems in science and engineering have a mathematical representation in the form of a system of Volterra type integral equations (VTIE). Dynamical systems represented through Delay-Differential Equations (DDE) are abundant in Chemical processes, Mechanical and Electrical systems etc.. Several researchers have studied the problem of optimal control and parameter determination in such systems.

The first results on optimal control of systems described by VTIE were due to Friedman (1964) and were followed by the work of Vinokurov (1969). Huang (1972) considered a general variational problem in such systems. Das (1967) considered the problem of optimal control for a hereditary process with delay, in presence of unknown parameters.

A Pontryagin's maximum principle for systems governed by DDE was derived by Kharatishvili (1961). Numerical techniques for the determination of optimal control in DDE have been developed by several authors. Eller et al. (1969) have presented a technique, that generates an open-loop control, in the case of

a quadratic cost functional. Slater and Wells (1972) have transformed the problem for a linear system, with the cost functional quadratic in control but not containing the state variables, into an equivalent non-delay problem. Sub-optimal controls have been obtained by Inoue et al. (1971), Jamshidi et al. (1972) and Malek-Zavarei (1980) using a MacLaurin series expansion of the control. Gracovetsky and Vidyasagar (1972,1973) and Malek-Zavarei (1980) have endeavoured to obtain sub-optimal controls, by solving a sequence of optimal control problems defined for ordinary differential equations.

In this work we consider some aspects of the above problems. A large portion of the work is devoted to the development of computational methods, for the determination of optimal control and parameters, in systems described by VTIE and DDE, in a deterministic framework.

Numerical algorithms, based on the quasilinearization technique, are developed for the determination of parameters in VTIE, minimizing a seminorm. The convergence proofs are given, under various system restrictions, for the two algorithms developed. As an application of these algorithms the problem of model reduction in dynamical systems is solved and various computational aspects are discussed. The method is illustrated by a numerical example. Then a set of necessary conditions is developed for optimal control of systems described by VTIE with parameters. This special treatment results in a compact set of conditions to be satisfied by the optimal control and parameters.

The problem of optimal control of systems described by DDE is then considered. An iterative scheme is proposed to determine a partly closed-loop and partly open-loop control for problems with quadratic cost functional. The convergence of the algorithm is established with certain system restrictions. As an illustration an optimal control policy, for a refining plant, to improve the system response is determined using the algorithm developed.

Finally the determination of optimal initial function and parameters in a system of DDE is discussed. The initial function is approximated to be contained in a finite dimensional subspace of the space of continuous functions and the problem is formulated as a parameter determination problem. Conditions for existence of a solution to the problem are discussed and a set of necessary conditions is developed. A steepest descent algorithm is proposed to determine a local solution of the problem. The method is illustrated by solving a model reduction problem.

CONTENTS

INTRODUCTION

Optimization theory had its origin in the pre-calculus era. Fermat's principle in optics and the isoperemetric problem, known as the problem of the Queen of Dido, were formulated before calculus was invented. The first results in optimization theory were systematized and brought together under the heading of the calculus of variations after the fundamental works of Euler (1707-1783). Euler and Lagrange formulated general constrained minimum problems and derived the appropriate multiplier rules in a formal manner. McShane [1] in his study of the Lagrange problem, constructed certain special variations which were shown to form a convex set. McShane's variations were used by Pontryagin et al. [2] in their proof of the maximum principle. The nature of this condition and the form of the optimal solutions are considerably different from the classical theorems of the Calculus of Variations and have occupied an important position in optimization theory. Starting with the basic results of Pontryagin et al. there have been notable contributions from a large number of scientists, which provide a good insight into the problems of optimization theory. Dubovitskii-Milyutin [3] found a necessary condition for an extremum in the form of an equation set down in the language of functional analysis. They were able to derive, as special cases of this condition, almost all previously known necessary

conditions. Gamkrelidze [4] formulated a general extremal problem in the theory of differential equations and derived necessary conditions, introducing the idea of a quasiconvex family of functions. Neustadt [5] formulated a general variational problem and derived necessary conditions, introducing the idea of first-order convex approximations - an extension of the cone of attainability introduced by Pontryagin et al. and the quasiconvexity introduced by Gamkrelidze. The Dubovitskii-Milyutin and the Gamkrelidze-Neustadt approaches differ basically in their treatment of the equality constraints. Further extensions and generalizations of the above results are being carried over till date. For some of the important contributions one can see the references [6,7,8,9].

Second order necessary conditions for optimization problems with constraints were developed by Dubovitskii-Milyutin [10]. Higher order necessary conditions in a very general setting (Ordered Topological Vector Space) were obtained by Hoffman et al. [11] by extending the concept of tangent cone used by Dubovitskii-Milyutin etc. Necessary and sufficient conditions for optimization problems in a Bannach Space with equality and inequality constraints were developed by Levitin et al. [12].

Economic and application necessities have given rise to the development of a vast literature for various specialized

problems. One can broadly view the general discipline of optimization theory as consisting of - Mathematical Programming, Optimal Control theory and Approximation theory.

Mathematical programming deals with the determination of extrema of smooth functions on closed domains with piecewise-smooth boundaries. First results in this direction were obtained by Kantorovich [13].

The problem of Optimal Control relates to finding the best (in some sense or other) among a given class of control functions so that a given process exhibits a desired behaviour.

The usual problem of approximation theory can be stated as follows :

Let (X,d) be a metric space and V be a subset of X. Let $v_0 \in X$ be a given point in X. It is required to find an element $v \in V$ such that $d(v_0, V)$ attains its minimum among all the points in V.

The following problem of optimal parameter determination can be thought of as a special problem of Approximation theory :

Determine a set of parameters (constants) associated with a given system so that certain system behaviours are best met. Here, by a system, we mean the mathematical representation of a physical object.

One comes across such problems in various areas of applied sciences and engineering. Some of these are listed below :

(i) Curve fitting :

A convenient and comprehensive representation of certain tabulated data is to represent it through a function relating the dependent and independent variables. One method of achieving this is to select a class of functions and choose one that best fits the data.

As an illustration; let the data consist of (y_i, x_i) ; $i = 1,2, \ldots, n$; y_i's are dependent variables and x_i's are independent variables. Let the class of functions be given by,

$$y(x) = \sum_{i=1}^{m} a_i \, |x - x_i| \sin \omega i x, \quad m < n$$

The values of the parameters, ω , a_i; $i = 1,2, \ldots, m$ are sought so as to satisfy the least square criterion, i.e.,

$$\sum_{i=1}^{m} (y(x_i) - y_i)^2 \text{ is a minimum.}$$

(ii) Model fitting :

The problem of model fitting consists of determining a mathematical model of a physical system. In this case the scientist is very-often aware of the laws governing the system under consideration. He can then describe the relationships among the observed quantities. For instance, the voltage current relationship for a linear inductor is given by,

$$i(t) = i(o)\, e^{-\frac{R}{L} t} + \frac{1}{L} \int_{o}^{t} e^{-\frac{R}{L}(t-s)} v(s)\, ds$$

The problem is to determine the values of R and L so as to best fit the measured data i(t), $t \in [0,T]$ corresponding to an applied voltage v(t), $t \in [0,T]$.

(iii) Model Reduction :

Model reduction comprises of obtaining an approximate representation of a complex system that makes it amenable to various system manipulations like design of controllers, analysis of system performance etc., maintaining permissible accuracy. A reduced model is most commonly used while dealing with large scale systems, e.g., power system analysis and design, process control and instrumentation etc.

As an illustration, consider the problem of designing regulators and stabilizers for a generator connected to a power system network. For this purpose the generator is represented by a detailed model while the rest of the network is represented by an equivalent approximate model of a small order.

A mathematical formulation of the model reduction problem can be stated as follows :

Let the system under consideration be represented by,

$$x(t) = f(t) + \int_0^t g(t,s, x(s), u(s))\,ds$$

$$y(t) = C(t, x(t), u(t)), \quad y \in R^m, \quad x \in R^n, \quad n \geq m$$

The problem is to determine the parameters a,b,c in the reduced system equation,

$$\bar{x}(t) = \bar{f}(t,a) + \int_0^t \bar{g}(t,s,\bar{x}(s), u(s), b)\, ds$$

$$\bar{y}(t) = \bar{C}(t,\bar{x}(t), u(t), c),\ \bar{y} \in R^m,\ \bar{x} \in R^q,\ n \geq q \geq m$$

such that $p(y-\bar{y})$ is a minimum, p is some given seminorm in $C_m[0,T]$; the space of all continuous functions in $[0,T]$ with range in R^m.

(iv) Input Signal Determination :

This problem is frequently encountered in measurement of signals. The output of a measuring instrument is a function of its system dynamics and the signal being measured. It may not always be possible to know the exact value of the signal being measured. In such a situation the problem is how to get a good approximant. One way of solving this could be as follows :

Let the measuring system dynamics be given by ;

$$x(t) = f(t, u(t)) + \int_0^t g(t,s,x(s),u(s))\, ds \qquad \text{(iv.1)}$$

$$y(t) = C(t,x(t))$$

y is the output of the measuring instrument for a signal u applied to its input terminals.

Consider a suitable set of n-linearly independent vectors $\{v_i\ ;\ i = 1,2, \ldots, n\} \in C_1[0,T]$. Let u be represented approximately by,

$$u(t) \approx \sum_{i=1}^{n} a_i \, v_i(t), \quad t \in [0,T] \qquad \text{(iv.2)}$$

Substituting (iv.2) in (iv.1) one gets,

$$\bar{x}(t) = f(t, \sum_{i=1}^{n} a_i \, v_i(t)) + \int_0^t g(t,s,\bar{x}(s), \sum_{i=1}^{n} a_i \, v_i(s)) \, ds$$

$$\bar{y}(t) = C(t,\bar{x}(t))$$

With this the problem reduces to the determination of the parameters a_i, $i = 1, \ldots, n$ such that $p(y-\bar{y})$ is minimum. Here p is a suitable seminorm in the appropriate space.

(v) Control Problems :

The optimal control problem can be treated within the parameter determination framework in special cases. In case the class of control functions depend on a finite number of parameters (i.e. it is assumed sufficiently smooth) the problem of optimal control reduces to the determination of parameters minimizing certain performance index associated with the problem.

As an illustration consider the following problem of programmed motion :

For the system in figure - 1 , the controller parameters $\underline{a}$ are to be chosen so that the system output $x(t)$ follows a prescribed trajectory $r(t)$, $t \in [0,T]$ as closely as possible.

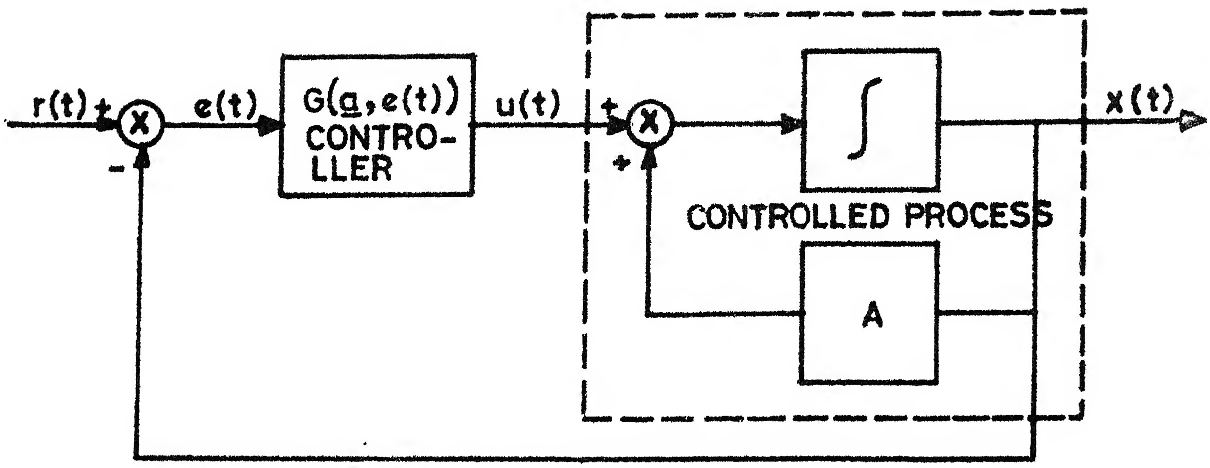

FIG.1 PROGRAMMED MOTION BLOCK DIAGRAM

A natural formulation of this problem would be to determine $\underline{a} \in R^m$, so as to minimize $||e||$ subject to the constraint,

$$e(t) = r(t) - \int_{o}^{t} \exp[A(t-s)]\ G(\underline{a}, e(s))\ ds$$

A - is the constant system matrix.

The problem of optimal control and parameter determination in systems described by Ordinary Differential Equations (ODE) has been understood and analysed to a fairly high degree of perfection.

A wide variety of systems in science and engineering have a mathematical representation in the form of a system of Volterra Type Integral Equations (VTIE). The above examples illustrate few such systems. A comprehensive list of systems represented by VTIE is given in [16]. Dynamical systems represented through Delay-Differential Equations (DDE) are abundant in Chemical processes, Mechanical and Electrical systems etc.

Kharatishvili [44] have developed a maximum principle for optimal control of systems described by DDE. Optimal control of systems governed by VTIE were first considered by Friedman[27]. Vinokurov [28] has considered the optimal control problem for VTIE with restricted phase coordinates. Huang [29] has considered more general variational problems and optimal control problems for a system of VTIE with and without restricted phase coordinates. He has derived a general maximum principle for such problems based on the multiplier rules developed by Neustadt [6].

Das [30] has considered the problem of optimal control of a class of VTIE with constant time lag and parameters.

One area of optimization theory which has received considerable attention, since the advent of high speed digital computers, is the numerical computation of the optimal solutions. A number of iterative schemes have been and are being proposed for the computation of the optimal solutions. Loosely speaking, all the available methods can be classified in to two catagories [17] : namely, (i) direct methods and (ii) indirect methods. In the direct methods the cost functional is minimized through an iterative scheme (e.g., the gradient method, penalty function approach etc.). Most of the indirect methods are essentially methods for the solution of the necessary conditions of optimality.

Gradient methods for determination of parameters and initial conditions for systems described by ODE are discussed by Bard [20], Junkins [21] etc. A quasilinearization technique for the determination of parameters and initial conditions, minimizing a seminorm by the direct method , is considered by Hoffman et al. [19] for systems described by ODE.

The numerical determination of the optimal control in systems governed by DDE is considered by several authors. Eller et al. [31] have presented a technique for solving the optimization problem with a quadratic cost functional. Slater and Wells [33] have considered the problem for linear time delay

systems with a quadratic cost functional involving only the control. A continuation or sensitivity approach is used by Inoue et al. [34] to determine sub-optimal controls for stationary systems with small delay in state. Jamshidi et al.[35] have expanded the control in a MacLaurin series and obtained the series coefficients from non-delay computations for a stationary linear system with delay. The result has been extended to multiple delays in state and control by Malek-Zavarei [36].

In this thesis we consider some aspects of the optimal control and parameter determination problems for systems described by VTIE and DDE. A large portion of the work is devoted to the computational aspects of the problems in a deterministic framework. The first three chapters deal with systems described by VTIE and the last two chapters with systems described by DDE.

In Chapter 1 numerical algorithms, based on quasilinearization technique, are developed for the parameter determination problem. The system equations are considered to be VTIE and the minimizing functional - a seminorm. This is a fairly general setting and can be applied to determine optimal parameters in a large class of practical problems. The convergence proofs are given for the two algorithms developed, under different restrictions on the system equations. It is seen that the first algorithm is quadratically convergent and requires much less computational effort in each iteration as compared to the second

algorithm but it imposes more constraints in the system equations, there by restricting it to a smaller class of problems.

In Chapter 2 the optimal model reduction problem is formulated to fit into the framework of the approximation problem considered in the previous section. The methods of solution for various system structures are discussed. This approach covers a large class of systems and simultaneously provides an easy way out from the difficulties arising in Wilson's [24] method for linear time invariant systems.

In Chapter 3 a set of necessary conditions are developed for optimal control of systems governed by VTIE with parameters. This special treatment results in a compact set of conditions and provides an insight in to this particular problem.

In Chapter 4 we consider the optimal control problem in systems described by DDE. An iterative scheme is presented to obtain a partly closed-loop and partly open-loop control for systems with delays in state variables. The works of Gracovetsky and Vidyasagar [38,39,40] and Malek-Zavarei [37] deal with similar problems. They have endeavoured to obtain sub-optimal controls of similar structure. The scheme presented in this section is easy to implement and gives an optimal control. The convergence proof utilizes a smallness condition. But, as it can be seen from the example considered, the method works even when the conditions are far from being satisfied.

Finally in Chapter 5 the problem of optimal initial function and parameter determination is considered for a system of DDE. To begin with, existance of a solution for the optimization problem is established. A set of necessary conditions is then developed in the form of a boundary value problem. A steepest descent algorithm is presented to obtain a solution using the set of necessary conditions.

CHAPTER I

DETERMINATION OF OPTIMAL PARAMETERS IN SYSTEMS DESCRIBED BY VOLTERRA TYPE INTEGRAL EQUATIONS

1. INTRODUCTION

The usual problem of Approximation Theory can be stated as follows :

Let (X,d) be a metric space and $V \subset X$ be a subset of X. Let $x_o \in X$ be a given point in X. It is required to find an element $v^* \in V \ni$

$$d(x_o,v^*) \leq d(x_o,v) \quad \forall v \in V.$$

The development of any computational method for the solution of this problem greatly depends on the nature of the space (X,d) and the set V. The work here is devoted to the development of numerical algorithms in the case X is restricted to the space of all continuous functions in $[0,1]$ with range in R^n and with the metric a seminorm. Further the set V is restricted to the set of solutions of a system of Volterra Type Integral Equations (VTIE) with parameters.

The algorithms developed are based on continuous linearization of the integral equations and the solution of the linear approximation problem. The basic idea in this is the quasi-linearization technique due to Bellman and Kalaba [18]. Hoffman

and Klostermair [19] have given the convergence proofs for these algorithms in case the set V is restricted to the set of solutions of a system of Ordinary Differential Equations with parameters. The extensions carried over here covers a much larger class of systems.

The algorithms can be conveniently used to solve a large number of problems in physics, engineering, bio-sciences etc., which require the determination of certain system parameters.

Section 2 contains the general problem formulation. The linearized problem is stated in Section 3. In section 4 we formulate the first algorithm (without λ-strategy) and give its convergence proofs. Section 5 contains the formulation of the second algorithm (with λ-strategy) and its convergence proof. In Section 6 we give methods of solution of the linearized problem in a few special cases. In the last section (Section 7) we present some discussions.

2. GENERAL PROBLEM FORMULATION

Let $\tilde{h} : [0,1] \times R^n \times R^m \times R^p \to R^n$, $q \in C_1[0,1]$ be given. Further let $T : C_n[0,1] \to C_1[0,1]$ be a linear operator. Here $C_r[0,1]$ denotes the space of all continuous functions on $[0,1]$ with range in R^r.

A solution $x^*(. \mid \mu^*, \beta^*)$ of the integral equation.

$$x(t) = \tilde{h}(t,x(t),\beta,\mu) = \tilde{f}(t,\mu) + \int_0^t \tilde{g}(t,s,x(s),\beta)\, ds \qquad (2.1)$$

is sought, so that $P(q - T x^*)$ is a minimum, under all the

solutions $x(.\,|\,\mu,\beta)$ of the equation (2.1). Here P is a semi-norm in $C_1[0,1]$.

In future the above problem will be assumed to have the following form :

$$x(t):=F(x,\mu)(t)=f(t,\mu)+\int_0^t g(t,s,x(s))\,ds$$

$F: \; C_n[0,1] \times R^p \to C_n[0,1]$ be a Frechet differentiable mapping and the derivative of F at (x,μ) will be denoted by $F'(x,\mu)$.

Let $A \subset R^p$ be convex. Define,

$B \subset C_n[0,1] \times R^p$ as $B:=C_n[0,1] \times A$

Problem (GP) :

Minimize $P(q-Tx)$

$(x,\mu) \in B \wedge x = F(x,\mu)$

3. LINEARIZED PROBLEM

Let $(\tilde{x},\tilde{\mu}) \in B$. Linearize F at $(\tilde{x},\tilde{\mu})$ and set,

$$F_{(\tilde{x},\tilde{\mu})}(x,\mu)=f(t,\tilde{\mu})+\frac{\partial f(t,\tilde{\mu})}{\partial \mu}(\mu-\tilde{\mu})+\int_0^t g(t,s,\tilde{x}(s))\,ds$$
$$+\int_0^t \frac{\partial g(t,s,\tilde{x}(s))}{\partial x}(x(s)-\tilde{x}(s))\,ds$$

The linearized problem at $(\tilde{x},\tilde{\mu})$ denoted by $(LP_{(\tilde{x},\tilde{\mu})})$ can be written as follows :

$(LP_{(\tilde{x},\tilde{\mu})})$ Minimize $P(q-Tx)$

$(x,\mu) \in B \wedge x = F_{(\tilde{x},\tilde{\mu})}(x,\mu)$

4. FORMULATION OF THE FIRST ALGORITHM AND CONVERGENCE PROOFS

For the formulation of the algorithm and proof of its convergence the following assumptions are used :

Assumption (V_1.a) : f is continuous and defined $\forall\, t \in R$

Assumption (V_1.b) : For each constant $\bar{K} > 0$ and each bounded $S \subset R^n$ $\exists$ a measurable function $K(t,s)$ $\ni$

$|g(t,s,x) - g(t,s,y)| \leq K(t,s)\,|x-y|$, wherever

$0 \leq s \leq t \leq \bar{K}$ and both x and y are in S.

For each $t \in [0,\bar{K}]$ the function $K(t,s)$ is in $L^1(0,t)$ as a function of s and $\lim\limits_{r \to 0} \int_t^{t+r} K(t+r,s)\, ds = 0$.

(This is obviously the case if $K(t,s)$ is integrable in the product space, by Fubini's Theorem).

(Here $|\cdot|$ denotes the Euclidean norm)

Further,

$$\int_0^1 \sup_{t \in [0,1]} K(t,s)\, ds \quad \text{exists.}$$

Lemma - 1 :

(i) $\forall\, (x,\mu) \in B \;\; \exists\, (\bar{x},\bar{\mu}) \in B \;\ni$

$$\bar{x} = F(\bar{x},\bar{\mu}) \quad \text{and} \quad \|\bar{x}-x\| \leq e^L \,\|x - F(x,\mu)\|$$

$$\text{where } L = \int_0^1 \sup_{t \in [0,1]} K(t,s)\, ds$$

(ii) $\forall\, (\tilde{x},\tilde{\mu}) \in B, \forall\, (x,\mu) \in B \;\; \exists\, (\bar{x},\bar{\mu}) \in B \;\ni$

$$\bar{x} = F_{(\tilde{x},\tilde{\mu})}(\bar{x},\bar{\mu}) \quad \text{and} \quad \|\bar{x}-x\| \leq e^L \,\|x - F_{(\tilde{x},\tilde{\mu})}(x,\mu)\|$$

Proof : (i)

Consider any $(x,\mu) \in B$.

$$\exists\ (\bar{x},\mu) \in B \ni$$

$$\|\bar{x}-x\|_t \leq \|f(t,\mu) + \int_0^t g(t,s,\bar{x}(s))\, ds - x(t)\|_t$$

(Here, $\|\cdot\|_t$ denotes norm in $C_r[0,t]$)

$$= \|f(t,\mu) + \int_0^t g(t,s,x(s))\, ds - x(t)$$

$$+ \int_0^t [g(t,s,\bar{x}(s)) - g(t,s,x(s))]\, ds\|_t$$

$$\leq \| x - F(x,\mu)\| + \int_0^t K(t,s)\|\bar{x}-x\|_s\, ds$$

Applying Gronwall's Lemma ;

$$\|\bar{x} - x\| \leq e^L \| x - F(x,\mu)\|$$

Proof : (ii)

The proof follows by replacing $F(\bar{x},\mu)$ by $F_{(\tilde{x},\tilde{\mu})}(\bar{x},\mu)$ in the above proof.

<u>Corollary : 1</u>

(i) $\forall\ K > e^L C$

$$\inf.\{P(q-Tx) + K\|x - F(x,\mu)\| : (x,\mu) \in B\}$$

$$= \inf.\{P(q-T\,x) : (x,\mu) \in B,\ x = F(x,\mu)\}$$

(ii) $\forall K > e^{L}C \quad \forall (\bar{x},\bar{\mu}) \in B$

$$P(q-T\,\bar{x}) + K\,||\,\bar{x} - F(\bar{x},\bar{\mu})\,||$$

$$= \inf.\,\{P(q-T\,x) + K\,||x - F(x,\mu)|| \;:\; (x,\mu) \in B\}$$

$\Longleftrightarrow (\bar{x},\bar{\mu})$ minimal solution of (GP)

(iii) Similar statements are valid for problem $(LP_{(\tilde{x},\tilde{\mu})})$; $(\tilde{x},\tilde{\mu}) \in B$; i.e.,

(a) $\forall K > e^{L}C$

$$\inf.\,\{P(q-T\,x) + K\,||x-F_{(\tilde{x},\tilde{\mu})}(x,\mu)|| \;:\; (x,\mu) \in B\}$$

$$= \inf.\,\{P(q-T\,x) \;:\; (x,\mu) \in B,\; x = F_{(\tilde{x},\tilde{\mu})}(x,\mu)\}$$

(b) $\forall K > e^{L}C \quad \forall (\bar{x},\bar{\mu}) \in B$

$$P(q-T\,\bar{x}) + K\,||\,\bar{x} - F_{(\tilde{x},\tilde{\mu})}(\bar{x},\bar{\mu})\,||$$

$$= \inf.\,\{P(q-T\,x) + K\,||x-F_{(\tilde{x},\tilde{\mu})}(x,\mu)|| \;:\; (x,\mu) \in B\}$$

$\Longleftrightarrow (\bar{x},\bar{\mu})$ minimal solution of $(LP_{(\tilde{x},\tilde{\mu})})$.

Here, $C := \sup.\,\{P(T\,x) \;:\; x \in C_n[0,1],\; ||x|| = 1\}$

<u>Proof</u> :

Proofs for (i) and (ii) are given here and (iii) follows in a similar manner.

(i) First consider the proof for '$\leq$'

$$\inf.\{P(q-Tx) + K\|x-F(x,\mu)\| : (x,\mu)\in B\}$$

$$\leqslant \inf.\{P(q-Tx) + K\|x-F(x,\mu)\| : (x,\mu)\in B,\ x = F(x,\mu)\}$$

$$\leqslant \inf.\{P(q-Tx) : (x,\mu)\in B,\ x = F(x,\mu)\}$$

Next we prove '$\geqslant$'

$$\inf.\{P(q-Tx) : (x,\mu)\in B,\ x = F(x,\mu)\}$$

$$\leqslant \inf.\{P(q-T\bar{x}) : (x,\mu)\in B,\ (\bar{x},\bar{\mu})\in B,\ \bar{x} = F(\bar{x},\bar{\mu}),$$
$$\|\bar{x} - x\| \leqslant e^{L}\|x - F(x,\mu)\|\}$$

$$\leqslant \inf.\{P(q-T\bar{x}) - P(T(x-\bar{x})) + C\|\bar{x}-x\| : (x,\mu)\in B,$$
$$(\bar{x},\bar{\mu})\in B,\ \bar{x} = F(\bar{x},\bar{\mu}),\ \|\bar{x}-x\| \leqslant e^{L}\|x-F(x,\mu)\|\}$$

$$\leqslant \inf.\{P(q-Tx) + C\|\bar{x}-x\| : (x,\mu)\in B,\ (\bar{x},\bar{\mu})\in B,$$
$$\bar{x} = F(\bar{x},\bar{\mu}),\ \|\bar{x}-x\| \leqslant e^{L}\|x-F(x,\mu)\|\}$$

$$\leqslant \inf.\{P(q-Tx) + C e^{L}\|x-F(x,\mu)\| : (x,\mu)\in B\}$$

$$\leqslant \inf.\{P(q-Tx) + K\|x-F(x,\mu)\| : (x,\mu)\in B\}$$

(ii) '$\Longrightarrow$'

$$\inf.\{P(q-Tx) + K\|x-F(x,\mu)\| : (x,\mu)\in B\}$$

$$= P(q-T\bar{x}) + K\|\bar{x} - F(\bar{x},\bar{\mu})\|$$

$$= P(q-T\bar{x}) + \bar{K}\|\bar{x} - F(\bar{x},\bar{\mu})\| + (K-\bar{K})\|\bar{x}-F(\bar{x},\bar{\mu})\|$$

(Here $\bar{K}$ is defined as $e^{L}C < \bar{K} < K$)

$$\geqslant \inf.\{P(q-Tx) : (x,\mu)\in B,\ x = F(x,\mu)\}$$
$$+ (K - \bar{K})\|\bar{x} - F(\bar{x},\bar{\mu})\|,$$

(by (i))

$= \inf. \{ P(q-T\ x) + K\ ||x-F(x,\mu)||: (x,\mu) \in B \} +$

$$+ (K-\overline{K})\ ||\ \overline{x} - F(\overline{x},\overline{\mu})\ ||$$

$\Longrightarrow \overline{x} = F(\overline{x},\overline{\mu})$

$\Longrightarrow P(q-T\ \overline{x}) = \inf. \{ P(q-T\ x) : (x,\mu) \in B,\ x = F(x,\mu) \}$

$\Longrightarrow (\overline{x},\overline{\mu})$ minimal solution of (GP)

4.1 ALGORITHM (A1) : [Iteration method without λ-strategy]

Start : Choose $(x^{(1)}, \mu^{(1)}) \in B$

Iteration step : $(x^{(i)}, \mu^{(i)})$ is already constructed.
Construct $(x^{(i+1)}, \mu^{(i+1)})$ as the minimal solution of the linear approximation problem $(LP_{(x^{(i)},\mu^{(i)})})$

It is assumed that such minimal solution exist.

For convergence proof, in addition to the assumptions (V1), the existence of a local minimal solution (x^*, μ^*) of (GP) with the following properties is essential :

<u>Assumption (V2)</u> : (x^*, μ^*) is a strongly unique local minimal solution of (GP) if the following holds :

Let $(x^*, \mu^*) \in B \wedge x^* = F(x^*,\mu^*)$ then

$$\exists \delta_1 > 0,\ K_1 > 0 \ni (x,\mu) \in B \wedge ||x-x^*|| \leq \delta_1 \wedge x = F(x,\mu) \Longrightarrow$$

$$P(q-T\ x) - P(q-T\ x^*) \geq K_1\ ||\ x - x^*\ ||$$

Lemma 2 :

Assumption (V2) $\Longrightarrow \forall (x,\mu) \in B \wedge \|x-x^*\| \leq \delta_1/2$

$$P(q-T\,x) - P(q-T\,x^*) + K_2\ \|x - F(x,\mu)\| \geq K_1\ \|x-x^*\|$$

where $K_2 = (C + K_1)\ e^L$.

Proof :

By Lemma (1·i)

$\exists\ (\bar{x},\bar{\mu}) \in B$ with $\bar{x} = F(\bar{x},\bar{\mu})$ and $\|\bar{x}-x\| \leq e^L \|x-F(x,\mu)\|$

Let $\|\bar{x}-x^*\| \leq \delta_1$. Then by Assumption (V2),

$$\begin{aligned}
K_1\ \|x-x^*\| &\leq K_1(\|x-\bar{x}\| + \|\bar{x}-x^*\|)\\
&\leq K_1\|x-\bar{x}\| + P(q-T\,\bar{x}) - P(q-T\,x^*)\\
&\leq K_1 e^L \|x-F(x,\mu)\| + P(q-T\,x) + P(T(x-\bar{x})) - P(q-T\,x^*)\\
&\leq K_1 e^L \|x-F(x,\mu)\| + P(q-T\,x) - P(q-T\,x^*) + C\ \|x-\bar{x}\|\\
&\leq (K_1+C)\ e^L \|x-F(x,\mu)\| + P(q-T\,x) - P(q-T\,x^*)\\
&= K_2\|x-F(x,\mu)\| + P(q-T\,x) - P(q-T\,x^*)
\end{aligned}$$

Suppose $\|\bar{x} - x^*\| > \delta_1$

$$\begin{aligned}
K_1\ \|x - x^*\| \leq K_1\ \|x-\bar{x}\| &= K_1\|x-\bar{x}\| + P(q-T\,x^*) - P(q-T\,x^*)\\
&\leq K_1\ \|x-\bar{x}\| + P(q-T(x^*-x+x)) - P(q-T\,x^*)\\
&\leq K_1\ \|x-\bar{x}\| + P(q-T\,x) - P(q-T\,x^*) + P(T(x-x^*))\\
&\leq K_1\|x-\bar{x}\| + P(q-T\,x) - P(q-T\,x^*) + C\,\|x-x^*\|\\
&\leq (K_1+C)e^L\|x-F(x,\mu)\| + P(q-T\,x) - P(q-T\,x^*)\\
&= K_2\ \|x-F(x,\mu)\| + P(q-T\,x) - P(q-T\,x^*)
\end{aligned}$$

Lemma 3 :

Assumption (V2) and convexity of B $\Longrightarrow \forall (x,\mu) \in B$,

$$P(q-Tx) - P(q-Tx^*) + K_2 \, || x - F_{(x^*,\mu^*)}(x,\mu) || \geq K_1 \, || x - x^* ||$$

Proof :

Let $x_\lambda := \lambda x + (1-\lambda)x^* = x^* + \lambda(x-x^*)$

and $\mu_\lambda := \mu^* + \lambda(\mu-\mu^*)$

For sufficiently small $\lambda > 0$ Lemma-2 and convexity of B $\Longrightarrow$

$$P(q-Tx_\lambda) - P(q-Tx^*) + K_2 \, || x_\lambda - F(x_\lambda,\mu_\lambda) || \geq K_1 \, || x_\lambda - x^* ||$$

$$\Longrightarrow \lambda P(q-Tx) - \lambda P(q-Tx^*) + K_2 \, || x_\lambda - F(x_\lambda,\mu_\lambda) || \geq \lambda K_1 \, || x - x^* ||$$

$$\Longrightarrow P(q-Tx) - P(q-Tx^*) + K_2 \, \left|\left| x - F(x^*,\mu^*) - \frac{F(x_\lambda,\mu_\lambda) - F(x^*,\mu^*)}{\lambda} \right|\right| \geq K_1 \, || x - x^* ||$$

$$\Longrightarrow P(q-Tx) - P(q-Tx^*) + K_2 \, \left|\left| x - F(x^*,\mu^*) - \frac{\partial f(\mu^*)}{\partial \mu}(\mu-\mu^*) - \int_0^t \frac{\partial g}{\partial x}(t,s,x^*(s))(x-x^*)_s \, ds \right|\right| \geq K_1 \, || x - x^* ||$$

$$\Longrightarrow P(q-Tx) - P(q-Tx^*) + K_2 \, || x - F_{(x^*,\mu^*)}(x,\mu) || \geq K_1 \, || x - x^* ||$$

Assumption (V3.a) : f is linear in μ.

Assumption (V3.b) : Let $\exists \; \delta_2 > 0$, a measurable function $\hat{K}(t,s) \ni$

$$\left| \frac{\partial g(t,s,\bar{x}(s))}{\partial x} - \frac{\partial g(t,s,x(s))}{\partial x} \right| \leq \hat{K}(t,s) \; |\bar{x} - x|_s$$

$\forall (\bar{x},\bar{\mu}) \in B \;\wedge\; (x,\mu) \in B$, wherever $|| \bar{x} - x^* || \leq \delta_2 \wedge || x - x^* || \leq \delta_2$.

Lemma-4 :

From Assumptions (V3) and convexity of B it follows that

$\forall\, (\bar{x},\bar{\mu}) \in B,\ (x,\mu) \in B$;

$$\| \bar{x}-x^* \| \leq \delta_2 \ \wedge \ \| x-x^* \| \leq \delta_2 \implies$$

$$\| F(x,\mu) - F_{(\bar{x},\bar{\mu})}(x,\mu) \| \leq \tfrac{1}{2} L_1 \ \| x-\bar{x} \|^2$$

$$\text{where} \quad L_1 = \sup_{t \in [0,1]} \int_0^1 \hat{K}(t,s)\, ds$$

Proof :

Define $x_\lambda := \bar{x} + \lambda(x-\bar{x})$

$$(F(x,\mu) - F_{(\bar{x},\bar{\mu})}(x,\mu))(t) = f(t,\mu) + \int_0^t g(t,s,x(s))\, ds - f(t,\bar{\mu})$$

$$- \frac{\partial f}{\partial \mu}(t,\bar{\mu})\ (\mu-\bar{\mu}) - \int_0^t g(t,s,\bar{x}(s))\, ds$$

$$- \int_0^t \frac{\partial g}{\partial x}(t,s,\bar{x}(s))(x-\bar{x})_s\, ds$$

$$= \int_0^t [g(t,s,x(s)) - g(t,s,\bar{x}(s))]\, ds - \int_0^t \frac{\partial g}{\partial x}(t,s,\bar{x}(s))\ (x-\bar{x})_s\, ds$$

$$= \int_0^1 \int_0^t \frac{\partial g}{\partial x}(t,s,x_\lambda(s))\ (x-\bar{x})_s\, ds\, d\lambda - \int_0^t \frac{\partial g}{\partial x}(t,s,\bar{x}(s))\ (x-\bar{x})_s\, ds$$

$$= \int_0^1 \int_0^t \left[\frac{\partial g}{\partial x}(t,s,x_\lambda(s)) - \frac{\partial g}{\partial x}(t,s,\bar{x}(s)) \right] (x-\bar{x})_s\, ds\, d\lambda$$

$$||F(x,\mu) - F_{(\bar{x},\bar{\mu})}(x,\mu)|| \leqslant \sup_{t\in[0,1]} \int_0^1 \int_0^t \left| \frac{\partial g}{\partial x}(t,s,x_\lambda(s)) - \frac{\partial g}{\partial x}(t,s,\bar{x}(s)) \right| ds.\ d\lambda\, ||x-\bar{x}||$$

$$\leqslant \sup_{t\in[0,1]} \int_0^1 \int_0^1 \hat{K}(t,s)\ |x_\lambda - \bar{x}|_s\ ds\ d\lambda\ \ ||x-\bar{x}||$$

$$\leqslant \sup_{t\in[0,1]} \int_0^1 \lambda \int_0^1 \hat{K}(t,s)\ ds\ d\lambda\ \ ||x-\bar{x}||^2$$

$$\leqslant \tfrac{1}{2} L_1\ \ ||x-\bar{x}||^2$$

4.1.1 FIRST CONVERGENCE PROOF FOR ALGORITHM (A1) :

<u>Theorem-1</u> :

Let A be convex and Assumptions (V1) be satisfied. Let (x^*,μ^*) be a local minimal solution of (GP). Also let Assumptions (V2) and (V3) be valid. Then the algorithm converges locally quadratically to (x^*,μ^*), i.e.,

$$\exists\ \delta > 0,\ \ C_1 > 0\ \ \ni (x^{(1)},\ \mu^{(1)}) \in B\ \wedge$$

$$||x^{(i)} - x^*|| \leqslant \delta \Longrightarrow \forall\ i \in N,\ ||x^{(i+1)} - x^*|| \leqslant C_1\, ||x^{(i)} - x^*||^2$$

where $\{(x^{(i)},\mu^{(i)})\}_{i\in N}$ is a sequence obtained through Algorithm (A1) with $(x^{(1)},\mu^{(1)})$ as its initial member. Here, N is the set of natural numbers.

<u>Proof</u> :

Choose δ_o with $0 < \delta_o < \delta_2$ and $L_1 \cdot K_2 \cdot \delta_o < K_1$.

Let $(x,\mu) \in B$ with $||x-x^*|| \leqslant \delta_o$ and $(\bar{x},\bar{\mu})$ be minimal

solution of $(LP_{(x,\mu)})$. Then by Lemma-3

$$K_1\|\bar{x}-x^*\| \leq P(q-T\,\bar{x}) - P(q-T\,x^*) + K_2\|\bar{x} - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})\|$$

$$\leq P(q-T\,\bar{x}) - P(q-T\,x^*) + K_2\|\bar{x} - F_{(x,\mu)}(\bar{x},\bar{\mu})\| +$$

$$K_2\|F_{(x,\mu)}(\bar{x},\bar{\mu}) - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})\| \qquad (4.1.1.i)$$

Consider,

$$\|F_{(x,\mu)}(\bar{x},\bar{\mu}) - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})\| = \|F(x,\mu) + \frac{\partial F}{\partial \mu}(x,\mu)(\bar{\mu}-\mu)$$

$$+ \frac{\partial F}{\partial x}(x,\mu)\;(\bar{x}-x) - F(x^*,\mu^*) - \frac{\partial F}{\partial \mu}(x^*,\mu^*)(\bar{\mu}-\mu^*)$$

$$- \frac{\partial F}{\partial x}(x^*,\mu^*)(\bar{x}-x^*)\|$$

$$= \|[F(x,\mu) + \frac{\partial F}{\partial \mu}(x,\mu)(\mu^*-\mu) + \frac{\partial F}{\partial x}(x,\mu)(x^*-x) - F(x^*,\mu^*)]$$

$$- \frac{\partial F}{\partial x}(x,\mu)\;x^* + \frac{\partial F}{\partial x}(x^*,\mu^*)\;x^*$$

$$+ (\frac{\partial F}{\partial x}(x,\mu) - \frac{\partial F}{\partial x}(x^*,\mu^*))\;\bar{x}\|$$

$$\leq \|F(x^*,\mu^*) - F_{(x,\mu)}(x^*,\mu^*)\| + \|(\frac{\partial F}{\partial x}(x,\mu) - \frac{\partial F}{\partial x}(x^*,\mu^*))$$

$$(\bar{x} - x^*)\| \qquad (4.1.1.ii)$$

Consider,

$$P(q-T\,\bar{x}) - P(q-T\,x^*) + K_2\|\bar{x} - F_{(x,\mu)}(\bar{x},\bar{\mu})\|$$

By Corollary-1 (iii.b)

$$P(q-T\,\bar{x}) + K_2 \,||\,\bar{x}-F_{(x,\mu)}(\bar{x},\bar{\mu})\,|| \leqslant P(q-T\,x^*) + K_2\,||\,x^*-F_{(x,\mu)}(x^*,\mu^*)\,||$$

$$\Longrightarrow P(q-T\,\bar{x}) - P(q-T\,x^*) + K_2\,||\,\bar{x}-F_{(x,\mu)}(\bar{x},\bar{\mu})\,||$$

$$\leqslant K_2\,||\,x^*-F_{(x,\mu)}(x^*,\mu^*)\,|| \qquad (4.1.1.iii)$$

Substituting in (4.1.1.i) from (4.1.1.ii) and (4.1.1.iii)

$$K_1\,||\bar{x}-x^*\,|| \leqslant 2K_2\,||\,F(x^*,\mu^*)-F_{(x,\mu)}(x^*,\mu^*)\,|| +$$

$$K_2\,||\,(\frac{\partial F}{\partial x}(x,\mu) - \frac{\partial F}{\partial x}(x^*,\mu^*))(\bar{x}-x^*)\,||$$

$$\leqslant K_2\,L_1\,||x-x^*||^2 + K_2\,L_1\,||\,x-x^*\,||\;||\,\bar{x}-x^*\,||$$

$$\leqslant K_2\,L_1\,||\,x-x^*||^2 + K_2\,L_1\,\delta_o||\,\bar{x}-x^*\,||$$

$$\Longrightarrow ||\,\bar{x}-x^*\,|| \leqslant \frac{K_2\,L_1}{K_1-L_1K_2\delta_o}\,||x-x^*||^2 = C_1\,||\,x-x^*\,||^2$$

So the assertion follows for every δ with $o < \delta < \delta_o$ and $C_1\delta < 1$.

REMARK

In the above convergence proof the function $\tilde{f}(t,\mu)$ is assumed to be continuous in t and linear in μ. The linearity in μ is not necessary if the function $\tilde{f}(t,\mu)$ is differentiable in t. In such a case the problem can be handled by modifying $\tilde{g}(t,s,x(s)\beta)$ to $\bar{g}(t,s,x(s),\beta,\mu)$ $(:= \frac{\partial\tilde{f}}{\partial t}(t,\mu)+\tilde{g}(t,s,x(s),\beta))$ and $\tilde{f}(t,\mu)$ to $\bar{f}(\mu)$ $(:= \tilde{f}(o,\mu))$. With this the problem is solvable with the above algorithm provided the modified problem satisfies the required conditions.

4.2 EXTENSION OF CONVERGENCE OF ALGORITHM (A1) :

In the previous section the function $f(t,\mu)$ was restricted to be linear in μ. The uniqueness in the local minimal solution of (GP) was assumed with respect to the space $C_n[0,1]$. In this section the case where $f(t,\mu)$ is nonlinear in μ has been studied. For the convergence proof of Algorithm (A1) the uniqueness of the solution is assumed in the set $B(:= C_n[0,1] \times A)$. In this case the theorem possesses a local character with respect to μ as well.

For this purpose the essential assumptions and auxilary results are given below :

<u>Assumption</u> (V2.m) :

(x^*,μ^*) is a strongly unique local minimal solution of (GP) if the following holds :

$$(x^*,\mu^*) \in B,\ x^* = F(x^*,\mu^*) \text{ and } \exists\ \delta_1 > 0,$$

$$K_1 > 0 \ni (x,\mu) \in B \ \wedge\ (\|x-x^*\| + \|\mu-\mu^*\|) < \delta_1$$
$$\wedge\, x = F(x,\mu) \Longrightarrow$$

$$P(q-T\,x) - P(q-T\,x^*) \geq K_1(\|x-x^*\| + \|\mu-\mu^*\|)$$

<u>Lemma-5</u> :

Assumption (V2.m) $\Longrightarrow$

$$\{\forall (x,\mu) \in B \ \wedge\ (\|x-x^*\| + \|\mu-\mu^*\|) \leq \delta_1/2 \Longrightarrow$$

$$P(q-T\,x) - P(q-T\,x^*) + K_2\|x-F(x,\mu)\|$$
$$\geq K_1(\|x-x^*\| + \|\mu-\mu^*\|)\}$$

where

$$K_2 = (C + K_1)\, e^L$$

<u>Proof</u> :

By Lemma-1 (i)

$\exists\ (\bar{x},\bar{\mu}) \in B$ with $\bar{x} = F(\bar{x},\bar{\mu})$ and

$$\|\bar{x}-x\| \leq e^L \; \|x-F(x,\mu)\|$$

In case $(\|\bar{x}-x^*\| + \|\bar{\mu}-\mu^*\|) \leq \delta_1$

by Assumption (V2.m)

$$K_1(\|x-x^*\| + \|\mu-\mu^*\|) \leq K_1(\|x-\bar{x}\| + \|\bar{x}-x^*\| + \|\bar{\mu}-\mu^*\| + \|\mu-\bar{\mu}\|)$$

$$\leq K_1\, e^L \|x-F(x,\mu)\| + P(q-T\,\bar{x}) - P(q-T\,x^*)$$

(by construction in Lemma-1(i), $\mu = \bar{\mu}$)

$$\leq K_1\, e^L \|x-F(x,\mu)\| + C\|x-\bar{x}\| + P(q-T\,x) - P(q-T\,x^*)$$

$$\leq K_2\, \|x-F(x,\mu)\| + P(q-T\,x) - P(q-T\,x^*)$$

Now consider the case $(\|\bar{x}-x^*\| + \|\bar{\mu}-\mu^*\|) > \delta_1$

$$K_1(\|x-x^*\| + \|\mu-\mu^*\|) \leq K_1(\|\bar{x}-x^*\| + \|\bar{\mu}-\mu^*\| -$$
$$\|x-x^*\| - \|\mu-\mu^*\|)$$

$$\leq K_1(\|\bar{x}-x\| + \|\bar{\mu}-\mu\|)$$

$$= K_1\, \|\bar{x}-x\|, \text{ since } \bar{\mu} = \mu$$

$$= K_1\, \|\bar{x}-x\| + P(q-T\,x^*) - P(q-T\,x^*)$$

$$\leq K_1\, \|\bar{x}-x\| + P(q-T\,x) + C\,\|x-x^*\| - P(q-T\,x^*)$$

$$\leq K_1\, \|\bar{x}-x\| + C\,\|\bar{x}-x\| + P(q-T\,x) - P(q-T\,x^*)$$

$$\leqslant (K_1+C)\, e^L \, ||x-F(x,\mu)|| + P(q-T\,x) - P(q-T\,x^*)$$

$$= K_2\, ||x-F(x,\mu)|| + P(q-T\,x) - P(q-T\,x^*)$$

<u>Lemma-6</u> :

Assumption (V2.m) and convexity of B $\Longrightarrow \forall (x,\mu) \in B$

$$P(q-T\,x) - P(q-T\,x^*) + K_2\, ||x-F_{(x^*,\mu^*)}(x,\mu)|| \geqslant K_1(||x-x^*|| + ||\mu-\mu^*||)$$

<u>Proof</u> :

Define $x_\lambda \; := \; \lambda x + (1-\lambda)x^* = x^* + \lambda(x-x^*)$ and

$\mu_\lambda = \mu^* + \lambda(\mu-\mu^*)$

For sufficiently small $\lambda > 0$, Lemma-5 and convexity of B $\Longrightarrow$

$$P(q-T\,x_\lambda) - P(q-T\,x^*) + K_2\, ||x_\lambda - F(x_\lambda, \mu_\lambda)|| \geqslant K_1(||x_\lambda-x^*|| + ||\mu_\lambda - \mu^*||)$$

$$\Longrightarrow \lambda P(q-T\,x) - \lambda P(q-T\,x^*) + K_2||\lambda x - \lambda F(x^*,\mu^*) + F(x^*,\mu^*) - F(x_\lambda,\mu_\lambda)|| \geqslant K_1\, \lambda\, (||x-x^*|| + ||\mu-\mu^*||)$$

$$\Longrightarrow P(q-T\,x) - P(q-T\,x^*) + K_2\, \left|\left|x-F(x^*,\mu^*) - \frac{F(x_\lambda,\mu_\lambda) - F(x^*,\mu^*)}{\lambda}\right|\right| \geqslant K_1(||x-x^*|| + ||\mu-\mu^*||)$$

$$\Longrightarrow P(q-T\,x) - P(q-T\,x^*) + K_2||x-F_{(x^*,\mu^*)}(x,\mu)|| \geqslant K_1(||x-x^*|| + ||\mu-\mu^*||)$$

as $\lambda \to 0$.

<u>Assumption (V3.m)</u> :

(a) $\frac{\partial f}{\partial \mu}$ is local Lipschitz continuous in a neighbourhood of μ^* ; i.e.,

$$\exists\, \delta_2' > 0,\quad L_1' > 0 \quad \forall\ \mu_1, \mu_2 \in R^p \ni$$

$$\|\mu_1 - \mu^*\| \leqslant \delta_2' \ \wedge\ \|\mu_2 - \mu^*\| \leqslant \delta_2' \Longrightarrow$$

$$\left\| \frac{\partial f(t,\mu_1)}{\partial \mu} - \frac{\partial f(t,\mu_2)}{\partial \mu} \right\| \leqslant L_1' \ \|\mu_1 - \mu_2\|$$

(b) Same as (V3.b)

<u>Lemma-7</u> :

From (V3.m) and convexity of B it follows that,

$$(\bar{x},\bar{\mu}) \in B, (x,\mu) \in B,\ \|\bar{x} - x^*\| \leqslant \delta_2,\ \|\bar{\mu} - \mu^*\| \leqslant \delta_2',$$

$$\|x - x^*\| \leqslant \delta_2,\ \|\mu - \mu^*\| \leqslant \delta_2' \Longrightarrow$$

$$\|F_{(\bar{x},\bar{\mu})}(x,\mu) - F(x,\mu)\| \leqslant \tfrac{1}{2}\,\bar{L}\ (\|x - \bar{x}\|^2 + \|\mu - \bar{\mu}\|^2)$$

where

$$\bar{L} = \text{Max.}\ (L_1', L_1)$$

<u>Proof</u> :

Define $x_{\lambda_1} = \bar{x} + \lambda_1 (x - \bar{x})$ and

$$\mu_{\lambda_2} = \bar{\mu} + \lambda_2 (\mu - \bar{\mu})$$

$$\|F_{(\bar{x},\bar{\mu})}(x,\mu) - F(x,\mu)\|_t = \Big\| \int_0^1 f'(\mu_{\lambda_2})(\mu - \bar{\mu})\, d\lambda_2 -$$

$$f'(\bar{\mu})(\mu - \bar{\mu}) + \int_0^1 G'(x_{\lambda_1})(x - \bar{x})\, d\lambda_1 - G'(\bar{x})\ (x - \bar{x}) \Big\|_t$$

Here

$f'(\mu_{\lambda_2})$ represents $\left.\frac{\partial f}{\partial \mu}\right|_{\mu=\mu_{\lambda_2}}$ and

$G'(x_{\lambda_1})(x-\bar{x})(t)$ represents $\int_0^t \frac{\partial g}{\partial x}(t,s,x_{\lambda_1}(s))(x-\bar{x})_s\, ds$

$$\|F_{(\bar{x},\bar{\mu})}(x,\mu) - F(x,\mu)\|_t \leq \int_0^1 \|f'(\mu_{\lambda_2}) - f'(\bar{\mu})\|_t\, d\lambda_2 \,\|\mu-\bar{\mu}\|$$

$$+ \int_0^1 \|G'(x_{\lambda_1}) - G'(\bar{x})\|_t\, d\lambda_1 . \|x-\bar{x}\|$$

$$\leq \int_0^1 \lambda_2 L_1'\, d\lambda_2 \,\|\mu-\bar{\mu}\|^2 + \int_0^1 \lambda_1 L_1\, d\lambda_1 \|x-\bar{x}\|^2$$

$$\leq \tfrac{1}{2} L_1 \|x-\bar{x}\|^2 + \tfrac{1}{2} L_1' \|\mu-\bar{\mu}\|^2$$

$$\Longrightarrow \|F_{(\bar{x},\bar{\mu})}(x,\mu) - F(x,\mu)\| \leq \tfrac{1}{2}\bar{L}\,(\|x-\bar{x}\|^2 + \|\mu-\bar{\mu}\|^2)$$

4.2.1 SECOND CONVERGENCE PROOF FOR ALGORITHM (A1) :

<u>Theorem-2</u> :

Let A be convex and Assumption (V1) be satisfied. Let (x^*,μ^*) be a local minimal solution of (GP) and assumptions (V2.m) and (V3.m) are valid. Then the algorithm (A1) converges locally quadratically to (x^*,μ^*), i.e.,

$$\exists\, \delta > 0,\ C_1 > 0 \ni (x^{(1)},\mu^{(1)}) \in B \wedge (\|x^{(i)}-x^*\| + \|\mu^{(i)} - \mu^*\|) \leq \delta$$

$$\Longrightarrow (\|x^{(i+1)}-x^*\| + \|\mu^{(i+1)}-\mu^*\|) \leq C_1(\|x^{(i)}-x^*\|^2 + \|\mu^{(i)} - \mu^*\|^2)$$

$\forall\ i \in N$, where,

$\{(x^{(i)}, \mu^{(i)})\}_{i \in N}$ is the sequence obtained through Algorithm (A1) with $(x^{(1)}, \mu^{(1)})$ as starting point; N is the set of natural numbers.

<u>Proof</u> :

Choose δ_o with $0 < \delta_o \leq \min(\delta_2', \delta_2)$ and $\bar{L}\, K_2\, \delta_o < K_1$. Let $(x,\mu) \in B$ with $(\, \|x-x^*\| + \|\mu-\mu^*\|\,) \leq \delta_o$ and $(\bar{x},\bar{\mu})$ the minimal solution of $(LP_{(x,\mu)})$.

From Lemma-6,

$$K_1(\|\bar{x}-x^*\| + \|\bar{\mu}-\mu^*\|) \leq P(q-T\,\bar{x}) - P(q-T\,x^*)$$
$$+ K_2 \|\bar{x} - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})\|$$

$$\leq P(q-T\,\bar{x}) - P(q-T\,x^*) + K_2\|\bar{x}-F_{(x,\mu)}(\bar{x},\bar{\mu})\|$$
$$+ K_2\|F_{(x,\mu)}(\bar{x},\bar{\mu}) - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})\| \qquad (4.2.1.i)$$

Consider,

$$(F_{(x,\mu)}(\bar{x},\bar{\mu}) - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})) = F(x,\mu) + F'(x,\mu)\ ((\bar{x},\bar{\mu}) - (x,\mu))$$
$$- F(x^*,\mu^*) - F'(x^*,\mu^*)\ ((\bar{x},\bar{\mu}) - (x^*,\mu^*))$$
$$= F(x,\mu) + F'(x,\mu)((x^*,\mu^*) - (x,\mu)) - F(x^*,\mu^*)$$
$$- F'(x,\mu)(x^*,\mu^*) + F'(x,\mu)\ (\bar{x},\bar{\mu})$$
$$+ F'(x^*,\mu^*)\ ((x^*,\mu^*) - (\bar{x},\bar{\mu}))$$

$$= -\,(F(x^*;\mu^*) - F_{(x,\mu)}(x^*,\mu^*)) - (F'(x,\mu) - F'(x^*,\mu^*)).$$
$$((x^*,\mu^*) - (\bar{x},\bar{\mu}))$$

$$\Longrightarrow \;\; \|F_{(x,\mu)}(\bar{x},\bar{\mu}) - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})\| \leq \|F(x^*,\mu^*) - F_{(x,\mu)}(x^*,\mu^*)\|$$

$$+ \|(F'(x,\mu) - F'(x^*,\mu^*))((x^*,\mu^*) - (\bar{x},\bar{\mu}))\| \quad (4.2.1.ii)$$

By Corollary-1 (iii.b)

$$P(q-T\,\bar{x}) + K_2\,\|\bar{x}-F_{(x,\mu)}(\bar{x},\bar{\mu})\| \leq \inf.\{\,P(q-T\,y)$$

$$+ K_2\|\,y-F_{(x,\mu)}(x,\tilde{\mu})\,\| : (x,\tilde{\mu}) \in B\}$$

$$\leq P(q-T\,x^*) + K_2\|\,x^* - F_{(x,\mu)}(x^*,\mu^*)\,\| \quad (4.2.1.iii)$$

Substituting in (4.2.1.i) from (4.2.1.ii) and (4.2.1.iii),

$$K_1(\|\bar{x}-x^*\| + \|\bar{\mu}-\mu^*\|) \leq 2K_2\|\,F(x^*,\mu^*) - F_{(x,\mu)}(x^*,\mu^*)\|$$

$$+ K_2\,\|(F'(x,\mu) - F'(x^*,\mu^*))((x^*,\mu^*) - (\bar{x},\bar{\mu}))\|$$

$$\leq K_2\,\bar{L}\,(\|x-x^*\|^2 + \|\mu-\mu^*\|^2) + K_2\,L_1\,\|\,x-x^*\|\,\|\bar{x}-x^*\|$$

$$+ K_2\,L_1'\,\|\,\mu-\mu^*\,\|\,\|\bar{\mu}-\mu^*\|$$

$$\leq K_2\,\bar{L}\,(\|\,x-x^*\|^2 + \|\,\mu-\mu^*\|^2) + K_2\,\bar{L}\,\delta_o(\|\bar{x}-x^*\,\|+\|\,\bar{\mu}-\mu^*\|\,)$$

$$\Rightarrow \;\; \|\,\bar{x}-x^*\| + \|\bar{\mu}-\mu^*\| \leq \frac{K_2\,\bar{L}}{K_1 - K_2\,\bar{L}\,\delta_o}(\|x-x^*\|^2 + \|\,\mu-\mu^*\|^2)$$

$$= C_1\,(\|\,x-x^*\|^2 + \|\,\mu-\mu^*\|^2)$$

Thus the assertion follows for every δ, with $0 < \delta < \delta_o$ and $C_1 \delta < 1$. ✓

5. FORMULATION OF THE SECOND ALGORITHM AND CONVERGENCE PROOF

This section presents an algorithm with λ -strategy; that converges with essentially weaker assumptions.

For this purpose the necessary assumptions and definitions are given below :

<u>Assumption $(\overline{V1})$</u> : Same as Assumption (V1)

<u>Assumption $(\overline{V2})$</u> : $K \in R^+$ with $K > e^{L}C$

$$0 < \beta < 1$$

$$(\lambda_j)_{j \in N} \subset R^+, \quad \lambda_1 = 1, \quad \alpha: = \inf_{j \in N} \frac{\lambda_{j+1}}{\lambda_j} > 0$$

$$\lim_{j} \lambda_j = 0$$

The set A is convex and compact.

<u>Definition-1</u> :

$(\hat{x}, \hat{\mu}) \in B$ is called a stationary point of (GP) iff $(\hat{x}, \hat{\mu})$ is a minimal solution of $(LP_{(\hat{x}, \hat{\mu})})$.

<u>Definition-2</u> :

The penalty function $\varphi : B \to R$ is defined as,

$$\varphi(x, \mu): = P(q-T\, x) + K \; || x-F(x, \mu) ||$$

Lemma-8 :

Let $(x,\mu) \in B$ and $(\bar{x},\bar{\mu})$ be the solution of $(LP_{(x,\mu)})$. Then,

$$P(q-T\,\bar{x}) = P(q-T\,\bar{x}) + K\,||\,\bar{x} - F_{(x,\mu)}(\bar{x},\bar{\mu})\,||$$

$$\leqslant P(q-T\,x) + K\,||\,x-F(x,\mu)\,|| = \varphi(x,\mu)$$

Proof :

We prove by contradiction. Suppose the assertion is not true. Then,

$$P(q-T\,\bar{x}) > P(q-T\,x) + K\,||\,x-F(x,\mu)\,||$$

$$> P(q-T\,\tilde{x}) - C\,||\,x-\tilde{x}\,|| + K\,||\,x-F(x,\mu)\,||$$

$$> P(q-T\,\tilde{x}) - C\,e^{L}\,||\,x-F_{(x,\mu)}(x,\mu)\,|| + K\,||\,x-F(x,\mu)\,||$$

$$(\text{taking } (\tilde{x},\tilde{\mu}) \ni\!-\; ||\,x-\tilde{x}\,|| \leqslant e^{L}\,||\,x-F_{(x,\mu)}(x,\mu)\,||)$$

$$> P(q-T\,\tilde{x}) + (K - C\,e^{L})\,||\,x-F_{(x,\mu)}(x,\mu)\,||$$

$$\Longrightarrow P(q-T\,\bar{x}) > P(q-T\,\tilde{x}) \wedge \quad \tilde{x} = F_{(x,\mu)}(\tilde{x},\mu)$$

$\Longrightarrow P(q-T\,\bar{x})$ is not a minimal solution of $(LP_{(x,\mu)})$

$$\Longrightarrow P(q-T\,\bar{x}) \leqslant \varphi(x,\mu) \quad \checkmark$$

Corollary-2 :

In case min $P(q-T\,x)$ over all $(x,\mu) \in B$ such that $x = F(x,\mu)$ is unique and min. $P(q-T\,\bar{x})$ over all $(x,\mu) \in B$ with $\bar{x} = F_{(x,\mu)}(\bar{x},\bar{\mu})$ is unique, then both the minima must coincide.

5.1 ALGORITHM (A2) : [Iteration method with λ-strategy]

Start : Choose $(x^{(1)}, \mu^{(1)}) \in B$, $x^{(1)} \in C_n[0,1]$,

$$\| x^{(1)} \| \leq M \geq 0.$$

Iteration Step :

Suppose $(x^{(i)}, \mu^{(i)})$ is already constructed. Determine $(\bar{x}^{(i)}, \bar{\mu}^{(i)})$ as the minimal solution of the linearized approximation problem $(LP_{(x^{(i)},\mu^{(i)})})$.

In case $P(q-T\,\bar{x}^{(i)}) = \varphi(x^{(i)}, \mu^{(i)})$; $(x^{(i)}, \mu^{(i)})$ is a stationary point of (GP) and the iteration will be stopped. Otherwise,

$$\varphi(x^{(i)}, \mu^{(i)}) > P(q-T\,\bar{x}^{(i)})$$

Take ;

$$h^{(i)} = \bar{x}^{(i)} - x^{(i)}$$

$$l^{(i)} = \bar{\mu}^{(i)} - \mu^{(i)}$$

Determine the smallest index $j := j^{(i)} \in N$ with

$$\varphi(x^{(i)} + \lambda_j h^{(i)}, \mu^{(i)} + \lambda_j l^{(i)}) \leq \varphi(x^{(i)}, \mu^{(i)}) - \beta\, \lambda_j(\varphi(x^{(i)}, \mu^{(i)}) - P(q-T\,\bar{x}^{(i)}))$$

Set,

$$x^{(i+1)} = x^{(i)} + \lambda_j\, h^{(i)}$$

$$\mu^{(i+1)} = \mu^{(i)} + \lambda_j\, l^{(i)}$$

A few further assumptions and definitions, essential for the convergence proof are given below :

Assumption $(\overline{V3})$: f is linear in μ.

Definition-3 :

The linearized penalty function, $\varphi_{(\hat{x},\hat{\mu})}: B \to R$, $(\hat{x},\hat{\mu}) \in B$ is defined as ;

$$\varphi_{(\hat{x},\hat{\mu})}(x,\mu) := P(q-T\,x) + K \,\| x - F_{(\hat{x},\hat{\mu})}(x,\mu) \|$$

Lemma-9 :

Let $(x,\mu) \in B \cdot (\overline{x},\overline{\mu}) \in B$ be the solution of $(LP_{(x,\mu)})$ with $\varphi(x,\mu) > P(q-T\,\overline{x})$.

Then $\exists\ \delta > o \ni$ for $o < \lambda < \delta$,

$$\varphi(x+ \lambda(\overline{x}-x), \mu+ \lambda(\overline{\mu}-\mu) \leq \varphi(x,\mu) - \beta\lambda\,(\varphi(x,\mu) - P(q-T\,\overline{x}))$$

Proof :

φ possesses in $(x,\mu) \in B$ a right derivative $\varphi'(x,\mu)$ and the following holds ;

$$\varphi'(x,\mu)\ ((\overline{x},\overline{\mu}) - (x,\mu)) = \varphi'_{(x,\mu)}(x,\mu)\ ((\overline{x},\overline{\mu}) - (x,\mu))$$

$$\leq \varphi_{(x,\mu)}(\overline{x},\overline{\mu}) - \varphi(x,\mu) = P(q-T\,\overline{x}) - \varphi(x,\mu) < 0.$$

From here the assertion follows as $0 < \beta < 1$.

Remark :

From Lemma-9 one concludes : (x^*, μ^*) is a local minimal solution of (GP)

$\Longrightarrow$ (x^*, μ^*) is a stationary point of (GP)

$\Longrightarrow$ $x^* = F(x^*, \mu^*)$

5.1.1 CONVERGENCE PROOF FOR ALGORITHM (A2) :

Theorem-3 :

Let A be convex, compact and Assumptions $(\overline{V1})$, $(\overline{V2})$ and $(\overline{V3})$ be satisfied. Starting with an arbitrary $(x^{(1)}, \mu^{(1)}) \in B$ $(\| x^{(1)} \| \leq M \geq 0)$ the iteration method (A2) is carried out. In case the method terminates at the i^{th} iteration step $(x^{(i)}, \mu^{(i)})$ must be a stationary point of (GP). Otherwise the method provides a bounded sequence $\{(x^{(i)}, \mu^{(i)})\}_{i \in N}$ in B, which possesses a limit point.

Every limit point is a stationary point of (GP).

Proof :

(i) The sequence possesses a limit point in B :

By construction $\{\varphi(x^{(i)}, \mu^{(i)})\}_{i \in N}$ is a monotone decreasing sequence and therefore for some constant $\xi > 0$ $\forall i \in N$;

$$\| x^{(i)} - F(x^{(i)}, \mu^{(i)}) \| \leq \xi$$

It follows from here, $\forall\ i \in N$, $\forall\ t \in [0,1]$;

$$|x^{(i)}(t)| \leq \xi + |F(x^{(i)}, \mu^{(i)})(t)|$$

$$\leq \xi + \|F(o,\mu^{(i)})\| + \int_0^t K(t,s) \, |x^{(i)}(s)| \, ds$$

By Gronwall's Lemma one obtains a bound for $x^{(i)}$ and hence an uniform bound from boundedness of A.

Obviously $\{x^{(i)}\}_{i \in N}$ is uniformly bounded and equicontinuous. So by Ascoli's theorem, $\{(x^{(i)}, \mu^{(i)})\}_{i \in N}$ possesses a limit point in B, as A is compact and f is continuous.

(ii) Every limit point of $\{(x^{(i)}, \mu^{(i)})\}_{i \in N}$ in B is a stationary point of (GP) :

Let $(x^*, \mu^*) \in B$ be a limit point of $\{(x^{(i)}, \mu^{(i)})\}_{i \in N}$. We prove by contradiction.

Let (x^*, μ^*) be a non-stationary point. Then,

$$\varphi(x^*, \mu^*) > m := \inf. \{ P(q - T x) : (x, \mu) \in B, \; x = F_{(x^*, \mu^*)}(x, \mu) \}$$

Since φ is continuous and $\{\varphi(x^{(i)}, \mu^{(i)})\}_{i \in N}$ is a monotone decreasing sequence, the following holds;

$$\varphi(x^*, \mu^*) = \inf. \{\varphi(x^{(i)}, \mu^{(i)}) : i \in N\}$$

Now we shall show that $\exists \, i \in N \ni$

$$\varphi(x^{(i+1)}, \mu^{(i+1)}) < \varphi(x^*, \mu^*) \qquad (5.1.1.i)$$

Let $\eta := \varphi(x^*, \mu^*) - m > o$ and $(\bar{x}, \bar{\mu}) \in B$ with

$$\varphi_{(x^*, \mu^*)}(\bar{x}, \bar{\mu}) = m.$$

Let $(\bar{x}^{(i)}, \bar{\mu}^{(i)})$ be a minimal solution of $(LP_{(x^{(i)}, \mu^{(i)})})$.

Consequently,

$$\varphi(x^{(i)},\mu^{(i)}) - P(q-T\,\bar{x}^{(i)}) \geq \varphi(x^*,\mu^*) - \varphi_{(x^{(i)},\mu^{(i)})}(\bar{x},\bar{\mu})$$

Choosing $(x^{(i)}, \mu^{(i)})$ sufficiently close to (x^*,μ^*) ;

$$|\varphi_{(x^{(i)},\mu^{(i)})}(\bar{x},\bar{\mu}) - \varphi_{(x^*,\mu^*)}(\bar{x},\bar{\mu})| \leq K\,||F_{(x^{(i)},\mu^{(i)})}(\bar{x},\bar{\mu}) - F_{(x^*,\mu^*)}(\bar{x},\bar{\mu})||$$

$$\leq K||F(x^{(i)},\mu^{(i)}) - F_{(x^*,\mu^*)}(x^{(i)},\mu^{(i)})|| + K||\int_0^t (\frac{\partial g}{\partial x}(t,s,x^{(i)}_{(s)})$$

$$- \frac{\partial g}{\partial x}(t,s,x^*(s)))(\bar{x}-x^{(i)})_s\, ds||$$

$$\leq \eta/2$$

From the above estimates,

$$\varphi(x^{(i)},\mu^{(i)}) - P(q-T\,\bar{x}^{(i)}) \geq \varphi(x^*,\mu^*)$$

$$-(\varphi_{(x^*,\mu^*)}(\bar{x},\bar{\mu}) + \eta/2) \geq \eta/2$$

Let $R > 0$ be $\ni \forall\, i \in N$, $||x^{(i)}|| \leq R \wedge ||\bar{x}^{(i)}|| \leq R$.

We have $\forall\, (x,\mu) \in B \;\exists\, \delta \ni\; 0 < \delta < 4R$

$\wedge\; ||x-x^*|| \leq \delta$

$$\Longrightarrow\; ||F(x,\mu) - F_{(x^*,\mu^*)}(x,\mu)|| \leq \frac{(1-\beta)\eta}{8R}\; ||x-x^*||$$

Now we give an estimate for λ_j $(j = j^{(i)})$ below :

For $0 < \lambda < \frac{\delta}{4R} \leq 1$ because of,

$$\| x^{(i)} + \lambda h^{(i)} - x^* \| \leq \| x^{(i)} - x^* \| + \lambda \| h^{(i)} \| \leq \frac{\delta}{2} + \frac{\delta . 2R}{4R} = \delta$$

the following holds ;

$$\varphi(x^{(i)} + \lambda h^{(i)}, \mu^{(i)} + \lambda l^{(i)}) \leq \varphi_{(x^*, \mu^*)} (x^{(i)} + \lambda h^{(i)}, \mu^{(i)} + \lambda l^{(i)})$$

$$+ K \| F(x^{(i)} + \lambda h^{(i)}, \mu^{(i)} + \lambda l^{(i)})$$

$$- F_{(x^*, \mu^*)} (x^{(i)} + \lambda h^{(i)}, \mu^{(i)} + \lambda l^{(i)}) \|$$

$$\leq \varphi_{(x^*, \mu^*)} (x^{(i)} + \lambda h^{(i)}, \mu^{(i)} + \lambda l^{(i)}) + \frac{(1-\beta)\eta}{8R} (\frac{\alpha\delta}{2} + 2R\lambda)$$

(taking $\| x^{(i)} - x^* \| \leq \frac{\alpha\delta}{2}$)

$$\leq \varphi_{(x^{(i)}, \mu^{(i)})} (x^{(i)} + \lambda h^{(i)}, \mu^{(i)} + \lambda l^{(i)}) + \frac{(1-\beta)\eta}{4} (\frac{\alpha\delta}{4R} + \lambda)$$

$$\leq (1-\lambda) \varphi_{(x^{(i)}, \mu^{(i)})} (x^{(i)}, \mu^{(i)}) + \lambda \varphi_{(x^{(i)}, \mu^{(i)})} (\bar{x}^{(i)}, \bar{\mu}^{(i)}) + \frac{(1-\beta)\eta}{4} (\frac{\alpha\delta}{4R} + \lambda)$$

$$\leq \varphi(x^{(i)}, \mu^{(i)}) - \lambda\beta(\varphi(x^{(i)}, \mu^{(i)}) - P(q-T\bar{x}^{(i)}))$$
$$+ \frac{(1-\beta)\eta}{4} (\frac{\alpha\delta}{4R} + \lambda) - \frac{(1-\beta)\eta\lambda}{2}$$

$$\leq \varphi(x^{(i)}, \mu^{(i)}) - \lambda\beta(\varphi(x^{(i)}, \mu^{(i)}) - P(q-T\bar{x}^{(i)}))$$
$$+ \frac{(1-\beta)\eta}{4} (\frac{\alpha\delta}{4R} - \lambda)$$

i.e.

$$\varphi(x^{(i)} + \lambda h^{(i)}, \mu^{(i)} + \lambda l^{(i)}) \leq \varphi(x^{(i)}, \mu^{(i)})$$
$$- \lambda\beta(\varphi(x^{(i)}, \mu^{(i)}) - P(q-T\bar{x}^{(i)}))$$

In case $\frac{\alpha\delta}{4R} \leq \lambda < \frac{\delta}{4R}$

Since, $\inf\{ \frac{\lambda_{j+1}}{\lambda_j} : j \in N \} = \alpha > 0$,

there is a $j' \in N$ with $\frac{\alpha\delta}{4R} \leq \lambda_{j'} < \frac{\delta}{4R}$

By the definition of $j^{(i)}$ in the Algorithm (A2) it follows,

$j = j^{(i)} \leq j'$ and $\lambda_j \geq \lambda_{j'} > \frac{\alpha\delta}{4R}$

The inequality (5.1.1.i) is obtainable as,

$$\varphi(x^{(i+1)}, \mu^{(i+1)}) = \varphi(x^{(i)} + \lambda_j h^{(i)}, \mu^{(i)} + \lambda_j l^{(i)})$$

$$\leq \varphi(x^{(i)}, \mu^{(i)}) - \lambda_j \beta(\varphi(x^{(i)}, \mu^{(i)}) - P(q-T \bar{x}^{(i)}))$$

$$\leq \varphi(x^*, \mu^*) + \frac{\alpha\delta\beta\eta}{16R} - \frac{\alpha\delta}{4R} \cdot \beta \frac{\eta}{2} < \varphi(x^*, \mu^*)$$

REMARK :

In the above convergence proof $\tilde{f}$ is assumed to be linear in μ. The proof can be extended to the case where $\tilde{f}$ is not necessarily linear in μ but is differentiable with respect to μ.

For the convergence proof in this general case it is only required to replace $||x||$ by $(||x|| + ||\mu||)$ and $||x_1 - x_2||$ by $(||x_1 - x_2|| + ||\mu_1 - \mu_2||)$ in the above proof.

6. METHODS OF SOLVING THE LINEAR PROBLEM

The linearized problem arising in the iteration schemes above is as follows :

$(LP_{(\tilde{x},\tilde{\mu})})$:

$$\text{minimize } P(q-T\,x)$$

$$(x,\mu) \in B \wedge x = F_{(\tilde{x},\tilde{\mu})}(x,\mu)$$

where,

$$x(t) = F_{(\tilde{x},\tilde{\mu})}(x,\mu)(t) = f(t,\tilde{\mu}) + \frac{\partial f}{\partial \mu}(t,\tilde{\mu})\ (\mu-\tilde{\mu})$$

$$+ \int_0^t g(t,s,\tilde{x}(s))\ ds + \int_0^t \frac{\partial g}{\partial x}(t,s,\tilde{x}(s))(x-\tilde{x})_s\ ds \quad (6.i)$$

$B = C_n[0,1] \times A$, $A \subset R^p$, A is a convex subset.

For simplicity of notation we write (6.i) as,

$$x(t) = F_{(\tilde{x},\tilde{\mu})}(x,\mu)(t) = \underline{a}(t) + D(t)\mu + \int_0^t K(t,s)\ x(s)\ ds \quad (6.ii)$$

We consider below the solution methods for few special cases of this problem :

6.1 PROBLEM WITH KNOWN RESOLVENT KERNEL :

If the resolvent Kernel for the system equation (6.ii) is known apriori, x can be explicitely expressed as a function of μ. Let R be the resolvent Kernel. Then,

$$x(t) = \underline{a}(t) + D(t)\ \mu + \int_0^t R(t,s)\ (\underline{a}(s) + D(s)\mu)\ ds$$

$= \underline{b}(t) + H(t)\ \mu$, where $\underline{b}$ and H are known functions,

$\underline{b} : [0,1] \to R^n$, $H : [0,1] \to R^n \times R^p$

The problem reduces to,

$$\underset{\mu \in A}{\text{Minimize}}\quad p(q - T(\underline{b} + H\ \mu))$$

This problem can be solved using a suitable technique depending on the seminorm P.

6.2 PROBLEM WITH DEGENERATE KERNEL

In this case the problem is easily reduced to one with linear differential equation constraints and thus can be solved using standard techniques.

To do this let us consider the system equation (6.ii). $x(t)$ can be written as ;

$$x(t) = \underline{a}(t) + D(t)\mu + \sum_{m=1}^{M} G_m(t)\, Y_m(t) \qquad (6.2.i)$$

where,

$$K(t,s) = \sum_{m=1}^{M} G_m(t)\, H_m(s) \qquad (6.2.ii)$$

$$Y_m(t) = \int_0^t H_m(s)\, x(s)\, ds \qquad (6.2.iii)$$

Now differentiating (6.2.iii) and substituting from (6.2.i) we obtain,

$$\dot{Y}_m(t) = H_m(t) \cdot \underline{a}(t) + H_m(t)\, D(t)\, \mu + H_m(t) \left(\sum_{i=1}^{M} G_i(t) \cdot Y_i(t) \right) \qquad (6.2.iv)$$

With this the problem of minimization turns out to be equivalent to ;

$$\underset{\mu \in A}{\text{Minimize}} \quad P(q - T(\underline{a}(t) + D(t)\mu + \sum_{m=1}^{M} G_m(t)\, Y_m(t,\mu)))$$

subject to

$$\dot{Y}_i(t;\ \mu) = H_i(t)\ [\underline{a}(t) + D(t)\mu + \sum_{m=1}^{M} G_m(t)\ Y_m(t;\mu)]$$

6.3 PROBLEM WITH INTEGRAL COST FUNCTIONAL

We give a set of sufficiency conditions for the following problem :

<u>Problem (P)</u> : Minimize $J := \int_0^1 f_o(s(x),s)\ ds$ (6.3.i)

subject to the constraints,

$$x(t) = f(t,\mu) + \int_0^t K(t,s)\ x(s)\ ds \qquad (6.3.ii)$$

where f and f_o are convex and differentiable with respect to μ and x respectively.

<u>Theorem-4</u> :

The pair (x^*,μ^*) is a solution of the optimization problem (P) if it satisfies the following conditions:

(i) $\exists$ a row vector $\psi : [0,1] \to R^n \ni$

$$\psi(t) = f_{o_x}(x^*(t),\ t) + \int_t^1 \psi(s)\ K(s,t)\ ds,\ t \in [0,1]$$

and $$\int_0^1 \psi(t)\ f_\mu(t,\mu^*)\ dt = 0$$

(ii) x^* is a solution of the constraint equation (6.3.ii) corresponding to the parameter value μ^*, i.e.,

$$x^*(t) = f(t,\mu^*) + \int_0^t K(t,s)\ x^*(s)\ ds,\ t \in [0,1]$$

Proof :

For any (x,μ) satisfying equation (6.3.ii) and (x^*,μ^*) as above, we have,

$$\Delta J = J(x,\mu) - J(x^*,\mu^*) = \int_0^1 [f_o(x(s),s) - f_o(x^*(s),s)]\, ds$$

$$\geq \int_0^1 f_{o_x}(x^*(s),s)\,(x^*(s) - x(s))\, ds \quad \text{(by convexity assumption)}$$

$$= \int_0^1 \Big[\psi(u)\, x^*(u) - \int_u^1 \psi(s)\, K(s,u)\, ds\, x^*(u) - \psi(u)\, x(u)$$

$$+ \int_u^1 \psi(s)\, K(s,u)\, ds\, x(u)\Big]\, du$$

$$= \int_0^1 \psi(u)\,[f(u,\mu^*) - f(u,\mu)]\, du \leq \int_0^1 \psi(u)\, f_\mu(u,\mu^*)(\mu^*-\mu)\, du$$

(by convexity)

$$= \int_0^1 \psi(u)\, f_\mu(u,\mu^*)\, du(\mu^*-\mu) = 0$$

$\Longrightarrow (x^*, \mu^*)$ is an optimal pair.

7. CONCLUSION

In this chapter we have given convergence proofs of two algorithms for the numerical determination of parameters in a system of VTIE to minimize a given seminorm. Each iteration of the algorithm without λ-strategy requires much less computational efforts as compared to that with λ-strategy and is quadratic convergent. But the former imposes stronger conditions on the functions f and g for its convergence and the convergence is local.

No doubt a large class of physical systems can be represented by VTIE but it does not cover all. One must investigate the convergence of these algorithms for systems described by Fredholm Integral Equations. An obvious extension, if possible, could be to consider systems with $n(n > 1)$ - independent variables.

The problem as formulated in this work deals with deterministic systems. A large class of systems in practice being stochastic, an attempt to formulate the problem in a stochastic setting will be useful.

The applicability of these algorithms for sequential estimation problems (with appropriate modifications if necessary) could be studied, so that these may be used for on-line parameter estimation problems.

CHAPTER II

OPTIMAL REDUCTION OF DYNAMICAL SYSTEMS

1. INTRODUCTION

The general problem of model reduction is to choose a system model from a given class that approximates, in some sense, the input-output characteristics of a given system. In the literature several methods are available to reduce the order of a linear time invariant system. These methods can be broadly classified as eigen value retention techniques [22,25] and non-eigenvalue retention techniques [23,24].

Most of the eigenvalue retention techniques are based on the retention of the dominant eigenvalues of the original system in the reduced model although there are exceptions to this. In Marshal's [25] method the group of eigenvalues to be retained is chosen by minimizing an integral square error criterion.

Anderson [23] and Wilson [24] have proposed methods where the original system eigenvalues need not be retained. Anderson's method is based on the evaluation of the projection of a specified vector on to a linear subspace. It fits the output data from the original system to a simplified model by the least square method.

Wilson gives equations for the determination of the optimal reduced model of a LTIVS by minimizing a functional on

the error between the output vector of the reduced model and the original model. The equations proposed are a set of non-linear matrix equations which do not possess a unique solution. For the solution of this set of equations and determination of an optimal reduced model, numerical algorithms incorporating standard minimization routines, which require the structure of the reduced system matrix to be constrained in some manner, have been proposed by many authors. The algorithm proposed by Mishra and Wilson [26] is self-contained and eliminates the constraints in the reduced model structure. But convergence is not guaranteed for any general LTIVS. Moreover the conditions for convergence of the algorithm for MIMO systems with arbitrary eigenvalues in the reduced model is not ascertained.

Anderson's method can be applied to time-varying systems provided that time variations are not rapid. In case of non-linear systems a linear reduced system could be obtained giving a least mean square fit over the range of operation of the non-linear plant. Very little is available in the literature for reducing a general dynamical system to other than a LTIV system structure. We have delt with a model reduction problem where the reduced model is a Volterra type integral equation. The algorithms developed in Chapter 1 are utilized. As the algorithms require the input-output functions, the importance of the structure of the original system is minimal except for the computation of the output function. Of course in certain cases the

original system equations can be used to get a better matching in some sense and reduction of computational effort as discussed in the case of LTIVS considered in Section 4.

2. REDUCTION OF A GENERAL SYSTEM

Let S be a given general system. We intend to obtain a reduced model S' of a given structure so that we obtain the best match between the outputs y and y' of the two systems corresponding to a given input u(t), $t \in [t_o, t_1]$. Mathematically ;

Let $y(t)$, $t \in [t_o, t_1]$ be the output of the original system, corresponding to a given input u(t), $t \in [t_o, t_1]$. Let the reduced model be,

$$S': \quad x(t) = f(t,\mu) + \int_{t_o}^{t} g(t,s,x(s),\, u(s),\, \beta)\, ds$$

$$y'(t) = C\, x(t)$$

where, f and g are known functions of their arguments and C is a constant matrix. The problem is to determine the parameters (μ,β) so that certain seminorm $P(y-y')$ defined over the prescribed interval $[t_o, t_1]$ is minimum.

The problem statement is exactly in the form of the problem considered in Chapter 1 ; where the original system output is a known function of t, once the response of the original system is computed for the given input function. The algorithms of Chapter 1 can be directly applied provided the functions

describing the reduced system satisfy the required smoothness conditions.

3. REDUCTION OF LTIVS :

Let,

$$S : \quad \dot{z} = A z + B u , \qquad z \in R^n, \quad u \in R^r , \quad y \in R^p$$

$$y = C z \qquad z(o) = z_o, \quad p < n$$

be a given system.

Our interest is to construct a reduced order model S_R for S such that the integral square error (ISE) between the outputs of the original and reduced system is minimum. Let,

$$S_R : \quad \dot{\bar{z}} = \bar{A} \bar{z} + \bar{B} u , \qquad \bar{z} \in R^m, \quad u \in R^r, \quad \bar{y} \in R^p$$

$$\bar{y} = \bar{C} \bar{z} \qquad m < n$$

The problem is to find the system matrices $(\bar{A},\bar{B},\bar{C})$ and the initial condition vector $\bar{z}(o) = \bar{z}_o$ such that,

$$J := \int_o^\infty (y-\bar{y})' (y-\bar{y}) \, dt, \text{ is a minimum.}$$

For the algorithms described in Chapter 1 to be applicable, one requires the interval of integration to be finite and the reduced system output matrix to be known apriori. Depending on the system requirements $\bar{C}$ matrix can be chosen conveniently. The minimization of J can be achieved for all practical purposes taking a sufficiently large interval $[0,T]$ instead of (o, ∞). The choice of T is quite subjective. It should be chosen large

enough to accomodate the prominent system properties such as system stability, system frequencies of oscillations, nature of the input function etc.. The implementation of the algorithms may not be of much use for problems so stated with arbitrary measurable essentially bounded forcing functions because of large computational time requirement; as in such a case one will have to take a very large value of T to obtain a satisfactory model. Of course one can conveniently handle problems with smooth forcing functions; e.g., periodic functions, exponential functions or polynomial functions etc..

In the next section we will give a modified statement of the problem for stable LTIVS. This class of systems is most frequently encountered in practice. We will discuss the various aspects of the reduction problem in detail, in this case.

4. REDUCTION OF STABLE - LTIVS

Let,

$$\begin{aligned} \bar{S} : \dot{z} &= \bar{A}\, z + \bar{B}\, u \quad , \quad z \in R^n, \quad u \in R^r, \quad y \in R^p \\ y &= \bar{C}\, z \quad , \quad z(o) = z_o, \quad n > p \end{aligned} \tag{4.i}$$

be a given stable linear time-invariant system. We are interested in constructing a reduced order stable model ;

$$\begin{aligned} S_R : \dot{\hat{x}} &= \hat{A}\, \hat{x} + \hat{B}\, u \quad , \quad \hat{x} \in R^m \quad , \quad u \in R^r, \quad \hat{y} \in R^p \\ \hat{y} &= \hat{C}\, \hat{x} \qquad m < n \end{aligned} \tag{4.ii}$$

such that the following performance criteria are satisfied :

(i) In steady state the output of the reduced model and the original system match exactly for any measurable essentially bounded forcing function; i.e.,

$$\lim_{t \to \infty} [y(t) - \hat{y}(t)] = 0$$

(ii) For the transient performance we minimize the sum of the ISE in the initial condition response and an upper bound on the ratio of the ISE due to forcing function and the integral of the L_2-norm of the forcing function. (This choice of the transient performance gives a good approximation for any square integrable input function), i.e.,

$$J = J_o + J'$$

where

$$J_o = \int_0^T < (\bar{C}\, e^{\bar{A}t}\, z_o - \hat{C}\, e^{\hat{A}t}\, \hat{x}_o), (\bar{C}\, e^{\bar{A}t}\, z_o - \hat{C}\, e^{\hat{A}t}\, \hat{x}_o) > dt$$

$J' :=$ some upper bound on $||y-\hat{y}|| \ / \int_0^T ||u||\, dt, \quad \int_0^T ||u||\, dt \neq o$

$||\cdot||$ represents L_2-norm.

We can take for J,

$$J = J_o + \sum_{i=1}^{r} J_i$$

$$J_i = \int_0^T < (\bar{C}\, e^{\bar{A}t}\, \bar{B} - \hat{C}\, e^{\hat{A}t}\, \hat{B})e_i, (\bar{C}\, e^{\bar{A}t}\, \bar{B} - \hat{C}\, e^{\hat{A}t}\, \hat{B})e_i > dt$$

where, e_i represents the i^{th} natural basis vector.

J_i represents the ISE due to unit impulse input function i.e., $u(t) = \delta_i(o)$

For $\bar{S}$ and S_R we have,

$$y(t) = \bar{C}\, e^{\bar{A}t}\, z(o) + \int_o^t \bar{C}\, e^{\bar{A}(t-s)}\, \bar{B}\, u(s)\, ds$$

$$\hat{y}(t) = \hat{C}\, e^{\hat{A}t}\, \hat{x}(o) + \int_o^t \hat{C}\, e^{\hat{A}(t-s)}\, \hat{B}\, u(s)\, ds$$

In order that the steady state matching criterion is satisfied, we must have,

$$\lim_{t\to\infty} [\bar{C}\, e^{\bar{A}t}\, z(o) - \hat{C}\, e^{\hat{A}t}\, \hat{x}(o) + \int_o^t [\bar{C}\, e^{\bar{A}(t-s)}\, \bar{B} - \hat{C}\, e^{\hat{A}(t-s)}\hat{B}]u(s)ds] = \underline{0}.$$

and this is ensured for all admissible u if,

$$\lim_{t\to\infty} \int_o^t [\bar{C}\, e^{\bar{A}(t-s)}\bar{B} - \hat{C}\, e^{\hat{A}(t-s)}\, \hat{B}]\, ds = [0]$$

Or equivalently,

$$\hat{C}\, \hat{A}^{-1}\, \hat{B} = \bar{C}\, \bar{A}^{-1}\, \bar{B} = \bar{D} \tag{4.iii}$$

(4.iii) is a set of non-linear equations in the elements of $\hat{A}$ and $\hat{B}$; $\hat{C}$ being a previously selected matrix.

Let B be the general solution of the equation;

$$\hat{C}[X] = \bar{D} \tag{4.iv}$$

Comparing (4.iii) and (4.iv) we can assume $\hat{B}$ of the form;

$$\hat{B} = \hat{A}\, B \tag{4.v}$$

The elements of B are functions of $n_o.r$ independent parameters. Here n_o is the dimension of the null space of $\hat{C}$.

With this we have the reduced system as,

$$S_R' : \dot{\hat{x}} = \hat{A}(\hat{x} + B\, u)$$

$$\hat{y} = \hat{C}\, \hat{x} \qquad (4.vi)$$

To meet the transient specification we will use the algorithms discussed in Chapter 1.

For this purpose we augment the system equations (4.vi) by considering the parameters to be determined, as state variables.

Let, $x(i)$, $\hat{x}(i)$ denote the ith component of the vectors x, $\hat{x}$ respectively. Define,

$$x(i) := \hat{x}(i) \quad ; \quad i = 1, \ldots, m$$

$$x(i.m+j) := \hat{A}_{ij} \quad ; \quad i = 1, \ldots, m \; ; \; j = 1, \ldots, m$$

and the $n_o.r$ independent parameters in B be denoted by,

$$x(m^2+m+i) \;, \quad i = 1, \ldots, n_o.r$$

Then the elements of B, are obtained as

$$B_{ij} = h_{ij}(x) \; ; \quad i = 1, \ldots, m; \quad j = 1, \ldots, r$$

where, h_{ij} are functions of $x \in R^{(m^2+m+n_o.r)}$ determined by the equations (4.iv).

With this the modified system is obtained as,

$$S_m : \quad \dot{x} = F(x,u)$$

$$\hat{y} = C\, x \;, \quad C = [\hat{C} \; : \; 0]$$

where,

$$F_i(x,u) := \sum_{j=1}^{m} x(im+j)\,[x(j) + \sum_{k=1}^{r} h_{ik}(x)\cdot u_k] \;, \; i = 1,\ldots,m$$

$$F_i(x,u) := 0 \;; \quad i = m+1, \ldots, (m^2+m+\eta_o \cdot r) \qquad (4.\text{vii})$$

The linearized system can be obtained by simple manipulations as

$$L: \quad \dot{x} = F_{\bar{x}}(x) = \tilde{A}(\bar{x})\, x + b(\bar{x}) \qquad (4.\text{viii})$$

$$\hat{y} = C\, x$$

where

$$\left.\begin{array}{l}
b_i(\bar{x}) = -\displaystyle\sum_{k=im+1}^{(i+1)m} \bar{x}(k-im)\;\bar{x}(k) - \sum_{k=(m^2+m+1)}^{(m^2+m+n_o\cdot r)} \\
\qquad\qquad [\displaystyle\sum_{j=1}^{m} \bar{x}(im+j)(\sum_{l=1}^{r} \frac{\partial h_{il}(\bar{x})}{\partial x_k} u_l)]\;\bar{x}(k) \;; \\
\qquad\qquad i=1, \ldots, m \\
\qquad\qquad \text{if } n_o\cdot r \geq 1 \\
b_i(\bar{x}) = -\displaystyle\sum_{k=im+1}^{(i+1)m} \bar{x}(k-im)\;\bar{x}(k), \quad i = 1, \ldots, m \;; \\
\qquad\qquad\qquad \text{if } n_o\cdot r = 0 \\
b_i(\bar{x}) = 0 \quad m < i \leq (m^2+m+n_o\cdot r)
\end{array}\right\} \qquad (4.\text{ix})$$

$$\tilde{A}_{ij}(\bar{x}) = \bar{x}(im+j) \;, \quad j = 1, \ldots, m;\; i = 1, \ldots, m$$

$$\tilde{A}_{ij}(\bar{x}) = \bar{x}(j-im) + \sum_{l=1}^{r} h_{il}(\bar{x})\, u_l;\; im < j \leq (m^2+m+n_o\cdot r) \qquad (4.\text{x})$$

$$i = 1, \ldots, m$$

$\tilde{A}_{ij}(\bar{x}) = 0$ for all other elements of $\tilde{A}$.

The linearized problem now is to minimize J subject to the constraint equation (4.viii).

Consider the functional,

$$I' = \int_0^T (y-C\,x)'\,(y-C\,x)\,dt \tag{4.xi}$$

where x is a solution of (4.iii)

Let $\dot{\Phi} = \tilde{A}(\bar{x})\Phi$, $\Phi(o) = I$, the identity matrix.

Then,

$$x(t) = \Phi(t)x_o + \int_0^t \Phi(t)\,\Phi^{-1}(s)\;b(\bar{x}(s))\,ds$$

$$= \Phi(t)x_o + \gamma(t) \tag{4.xii}$$

where, $\gamma(t)$ is a solution of (4.viii) with $x(o) = \underline{0}$.

Substituting for (4.ii) in (4.xi) one obtains the equivalent problem of minimizing I' as follows ;

Find a set of initial conditions x_o^* for (4.viii) such that,

$$I = x_o'\,D\,x_o - 2\,x_o'\,E \quad \text{is a minimum.}$$

where,

$$D = \int_0^T \Phi(t)'\,C'\,C\,\Phi(t)\,dt\;;\; E = \int_0^T \Phi(t)'\,C'[y(t) - C\gamma(t)]\,dt \tag{4.xiii}$$

Consider the minimization of J_o ;

We have $y_o = \bar{C}\,z^o$, z^o a solution of $\bar{S}$ with $z(o) = z_o$ and $u = o$. The modified system $M_o(: \dot{x} = F^o(x,u))$ and the linearized

system $L_o(: \dot{x} = \tilde{A}^o(\bar{x}^o)\ x + b^o(\bar{x}^o))$ for this problem are obtained by substituting $u = o$ respectively in M and L derived above. The corresponding J_o is obtained by substituting for Φ and γ in (4.xiii) respectively Φ_o and γ_o as obtained from L_o and for y we substitute y_o. This gives, $J_o = x_o'\ D^o\ x_o - 2\ x_o'\ E^o$.

Consider the minimization of J_i, $i = 1, \ldots, r$. As in the case of J_o the modified systems M_i and L_i are obtained by substituting $u_j = 0$, $j = 1, \ldots, r$ $(j \neq i)$, $u_i = \delta(o)$ and $\underline{y_i}$ obtained with $z(o) = o$, $u_i = \delta(o)$, $u_j = o$, $j = 1, \ldots, r$ $(j \neq i)$ from the respective equations M, L and $\bar{S}$. Φ_i and $\underline{\gamma}_i$ are accordingly obtained as the state transition matrix and the appropriate solution of L_i. With these, $J_i = x_o'\ D^i\ x_o - 2\ x_o'\ E^i$.

Now summing up J_i's we obtain,

$$J = \sum_{i=0}^{r} J_i = x_o' \left[\sum_{i=0}^{r} D^i \right] x_o - 2\ x_o' \left[\sum_{i=0}^{r} E^i \right]$$

$$:= x_o'\ \bar{D}\ x_o - 2\ x_o'\ \bar{E}$$

The linearized problem can be stated as;

$$(LP_{\bar{x}}^{\ j}\ ;\ j = o, \ldots, r) : \quad \underset{x(o) \in S}{\text{minimize}}\ J = x_o'\ \bar{D}\ x_o - 2x_o'\ \bar{E}$$

The iteration scheme for the present reduction problem is as follows :

Start : Choose $x^{(o)} \in S$

Iteration Step :

Suppose $x^{(i)}(o)$ is already constructed. Construct $\bar{x}^{(i)}(o)$ as the minimal solution of $(LP_{\bar{x}}^{j}\ ,\ j = o, \ldots, r)^{(i)}$. In case

$$\sum_{j=o}^{r} ||\underline{y}_j - C\,\bar{x}^{j(i)}|| = \sum_{j=o}^{r} ||\underline{y}_j - C\,x^{j(i)}||$$

$$+ K\left(\sum_{j=o}^{r} ||\dot{x}^{j(i)} - F^j(x^{j(i)})||\right),$$

$x^{(i)}(o)$ is a solution of the reduction problem and the iteration will be stopped, otherwise put $h^{j(i)} = \bar{x}^{j(i)} - x^{j(i)}$, $j = o, \ldots, r$. Determine the smallest index $l := l^i \in N$ with,

$$\sum_{j=o}^{r} ||\underline{y}_j - C[x^{j(i)} + \lambda_l h^{j(i)}]|| + K\left(\sum_{j=o}^{r} ||\dot{x}^{j(i)} + \lambda_l\,\dot{h}^{j(i)}\right.$$

$$\left. - F^j(x^{j(i)} + \lambda_l\,h^{j(i)})||\right) \leq \left[\sum_{j=o}^{r} ||\underline{y}_j - C\,x^{j(i)}||\right.$$

$$\left. + K\left(\sum_{j=o}^{r} ||\dot{x}^{j(i)} - F^j(x^{j(i)})||\right]\right)(1-\beta\lambda_l) + \beta\lambda_l\left(\sum_{j=o}^{r} ||\underline{y}_j - C\,\bar{x}^{j(i)}||\right)$$

Set, $x^{j(i+1)} = x^{j(i)} + \lambda_l\,h^{j(i)}$, $j = o, \ldots, r$.

The above algorithm is for a fixed initial condition for the original system. This can be generalized to the case of all bounded initial condition vectors with a proper choice of J. The only change in the algorithm above is the index j ranges from $-(n-1)$ to r instead of 0 to r.

4.1 COMPUTATIONAL ASPECTS

For the implementation of the algorithm proposed above with

economization of computation time and to simultaneously obtain a reasonable accurate result one has to carefully choose the required subroutines.

Beyond this the various other factors those can contribute a significant reduction of computation time are discussed below :

(a) Choice of the interval length $[0,T]$:

One can considerably reduce the computation time by a judicious choice of the interval length $[0,T]$ of minimization; which is problem dependent. A choice of T approximately equal to the largest time constant of the system should be satisfactory in most cases as all the important transient characteristics of the system can be considered to be contained in this interval. It may be possible to choose a smaller interval in some cases. Care should be taken to retain various prominent system characteristics. The interval should be large enough so that the reduced system retains stability. The choice of a proper time interval reduces appreciable amount of computation time maintaining a good overall matching.

(b) Exploitation of the structure of the linearized system and the associated functional :

Consider the system L and the associated functional I'. The system matrix $\tilde{A}(\bar{x})$, $b(\bar{x})$, C and Φ can be partitioned as,

$$\tilde{A}(\bar{x}) = \begin{bmatrix} [\tilde{A}_{11}(\bar{x})]_{mxm} & [\tilde{A}_{12}(\bar{x})_{mx(m^2+n_o r)} \\ [0] & [0] \end{bmatrix}$$

$$b(\bar{x}) = \begin{bmatrix} \underline{b}_1(\bar{x})_{mx1} \\ \hline 0 \end{bmatrix}, \qquad C = [\hat{C} \ \vdots \ 0]$$

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix}$$

With the above partitioning of $\tilde{A}(\bar{x})$ and $b(\bar{x})$ one is required to solve only m differential equations numerically to obtain $(m^2 + m + n_o r)$ vector functions x and $\Upsilon(= (\frac{\Upsilon_1}{o}))$. For solving $\dot{\Phi} = \tilde{A}(\bar{x})\,\Phi$, $\Phi(o) = I$ one needs numerical solution of

$$\dot{\Phi}_{11} = \tilde{A}_{11}(\bar{x}) \quad \Phi_{11} \quad , \qquad \Phi_{11}(o) = I_{mxm}$$

and

$$\dot{\Phi}_{12} = \tilde{A}_{12}(\bar{x}) \quad \Phi_{12} + \tilde{A}_{12}(\bar{x}), \qquad \Phi_{12}(o) = [0]$$

The D and E matrices are obtained as,

$$D: = \begin{bmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{bmatrix} = \int_0^T \begin{bmatrix} \Phi_{11}'\,\hat{C}'\,\hat{C}\,\Phi_{11} & \Phi_{11}\,\hat{C}'\,\hat{C}\,\Phi_{12} \\ \Phi_{12}'\,\hat{C}'\,\hat{C}\,\Phi_{11} & \Phi_{12}'\,\hat{C}'\,\hat{C}\,\Phi_{12} \end{bmatrix} dt$$

$$E = \int_0^T \begin{bmatrix} \Phi_{11}'\,\hat{C}'\,[y - \hat{C}\,\Upsilon_1] \\ \Phi_{12}'\,\hat{C}'\,[y - \hat{C}\,\Upsilon_1] \end{bmatrix} dt = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$$

With the system variables represented in the above partitioned form, one can obtain a considerable reduction in computation time.

(c) Specialized problems :

A reduction in computation time is very often possible by exploiting the particular problem statement, i.e., by specializing the algorithm or a part of it for a particular problem one is interested to solve. We discuss below some important cases :

(i) Fixed initial condition for reduced model :

In this case one needs to compute only D_{22} and E_2 in D and E respectively. This avoides the solution of the m^2 - differential equations $\dot{\Phi}_{11} = \tilde{A}_{11}(\bar{x})\ \Phi_{11}$. Also the minimization of the linearized problem is to be carried out for the functional,

$$J = x'\ D_{22}\ x - 2\ E_2'\ x\ , \quad x \in S \subset R^{m^2+n_0 \cdot r}$$

(ii) Reduction for a particular input function with fixed initial condition for the reduced model :

In this case a considerable amount in computation time can be saved for a multi-input system by taking the minimizing functional as the ISE and making use of the general algorithm.

(iii) Reduced system matrix with specified structure :

If the structure of the reduced system matrix is specified, it may be possible to reduce the order of the modified system and subsequently the associated systems. This may reduce to some extent the computation time requirement. In certain

problems some elements of the reduced system matrices may be specified. In this case the order of the modified system and the associed systems definitely reduce.

4.2 NUMERICAL EXAMPLE

As an illustration we consider the system equations in [25] :

$$\bar{S} : \quad \dot{z} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -0.07 & -0.66 & -1.99 & -2.4 \end{bmatrix} z + \begin{bmatrix} 1.0 \\ -1.06 \\ 1.147 \\ -1.2544 \end{bmatrix} u$$

$$y = [1 \ 0 \ 0 \ 0] \ z, \qquad z(o) = \underline{0}$$

Our interest is to find a reduced model of second order satisfying the performance criteria discussed in this section. Further let the problem require the restriction of the structures of the reduced system matrices to the following form :

$$S_R : \quad \dot{\hat{x}} = \begin{bmatrix} 0 & 1 \\ \hat{a}_1 & \hat{a}_2 \end{bmatrix} \hat{x} + \begin{bmatrix} 0 \\ \hat{b} \end{bmatrix} u$$

$$\hat{y} = [1 \ 0] \ \hat{x}$$

and let the reduced model initial condition be given to be, $\hat{x}(o) = \underline{0}$.

From the steady state specification the system S_R' is obtained as,

$$S_R' : \quad \dot{\hat{x}} = \begin{bmatrix} 0 & 1 \\ \hat{a}_1 & \hat{a}_2 \end{bmatrix} (\hat{x} + \begin{bmatrix} d \\ o \end{bmatrix} u)$$

$$\hat{y} = [1,0] \ \hat{x} , \qquad \hat{x}(o) = \underline{0}$$

where, $d = C\ A^{-1}\ b$; C, A and b are the original system quantities.

Thus d is a known constant and the parameters to be determined are $\hat{a}_1$ and $\hat{a}_2$.

Let,

$$x_1 = \hat{x}_1, \quad x_2 = \hat{x}_2, \quad x_3 = \hat{a}_1, \quad x_4 = \hat{a}_2$$

with this the modified system equations are obtained as,

$$S_m: \quad \begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3\ x_1 + x_1\ x_2 + d\ x_3\ u \\ \dot{x}_3 &= 0 \\ \dot{x}_4 &= 0 \\ \hat{y} &= [1\ 0\ 0\ 0\ ]x, \quad x_1(o) = o, \quad x_2(o) = o \end{aligned}$$

Linearizing this at a point $\bar{x} \in C_4[0,T]$, we obtain,

$$L: \dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \bar{x}_3 & \bar{x}_4 & \bar{x}_1+du & \bar{x}_2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ -\bar{x}_3\ \bar{x}_1 - \bar{x}_2\ \bar{x}_4 \\ 0 \\ 0 \end{bmatrix}$$

The functional to be minimized is,

$$J = X'\ D_{22}X - 2E_2'X \quad , \qquad X \in R^2$$

where

$$D_{22} = \int_0^T \Phi_{12}' \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Phi_{12}\ dt$$

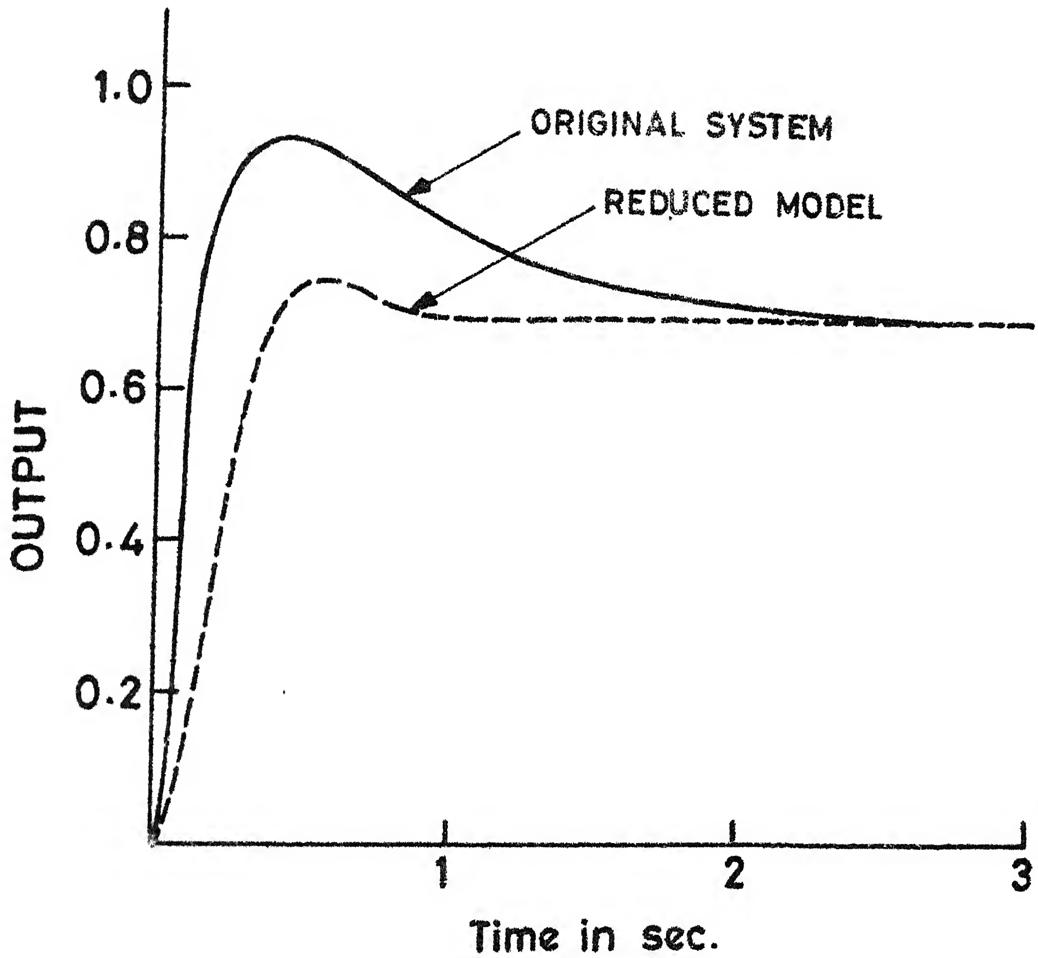

FIG. II.1 SYSTEM RESPONSE FOR UNIT STEP INPUT

$$E_2 = \int_0^T \Phi_{12}' \begin{bmatrix} 1 \\ 0 \end{bmatrix} [y - [1 \;\; 0] \, \underline{\gamma}_1] \, dt$$

$$\dot{\Phi}_{12} = \begin{bmatrix} 0 & 1 \\ \bar{x}_3 & \bar{x}_4 \end{bmatrix} \Phi_{12} + \begin{bmatrix} 0 & 0 \\ \bar{x}_4 + du & \bar{x}_2 \end{bmatrix}, \Phi_{12}(o) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\dot{\underline{\gamma}}_1 = \begin{bmatrix} 0 & 1 \\ \bar{x}_3 & \bar{x}_4 \end{bmatrix} \underline{\gamma}_1 + \begin{bmatrix} 0 \\ -\bar{x}_3 \, \bar{x}_1 - \bar{x}_2 \, \bar{x}_4 \end{bmatrix}, \quad \underline{\gamma}_1(o) = 0$$

For the given system the smallest eigenvalue is -0.2. We choose $T = \frac{1}{0.2} = 5$ seconds.

The problem was solved, with initial guess (-1.0, -1.0) for $(x_3(o), x_4(o))$.

The reduced system was obtained as,

$$S_R : \dot{\hat{x}} = \begin{bmatrix} 0 & 1 \\ -0.58235 & -1 \end{bmatrix} \hat{x} + \begin{bmatrix} 0.0 \\ 0.40764 \end{bmatrix} u$$

$$\hat{y} = [1, 0] \, \hat{x}$$

The system response for unit step input is shown in figure (II.1).

5. CONCLUSION

A need for model reduction arises when one intends to derive simplified control and estimation strategies. For simulation and control of complex processes one has to go for a reduced model; be it from economic point of view or to make the important system characteristics amneable.

The reduction algorithm discussed in this section can be suggested as an alternative scheme for various special problems. It can be conveniently used for the general problem, with the reduced system dynamics being restricted to the class of Volterra type integral equations. For the reduction of stable LTIVS, this algorithm gives a better approximation than that obtained by eigenvalue retention techniques as in the later case the choice of the parameters of the reduced system matrices is more restrictive.

CHAPTER III

NECESSARY CONDITIONS FOR OPTIMAL CONTROL OF SYSTEMS DESCRIBED BY VOLTERRA TYPE INTEGRAL EQUATIONS WITH PARAMETERS

1. INTRODUCTION

The problem of optimal control of systems described by Volterra Type Integral Equations (VTIE) was first considered by Friedman [27]. He considered the system equations with convolution kernel and derived a maximum principle for the case of unconstrained state variables. The cost functional had a structure similar to the system equations. Vinokurov [28] obtained maximum principle for a system described by general VTIE with restricted state variables. Huang [29] has considered the general variational problem for a system of VTIE. He generalizes the concept of quasiconvexity due to Gamkrelidze [4] and derives the general maximum principle based on the multiplier rules developed by Neustadt [6]. The problem of optimal control of systems governed by ordinary differential equations with parameters was considered by Boltyanskii [2]. Das [30] has considered the problem defined for VTIE with constant time-delay and with parameters. The system equations in his formulation are restricted to that with convolution Kernel.

In this chapter we consider the optimal control of systems governed by VTIE with parameters and a general kernel. An

integral cost functional of a similar structure is considered and a set of necessary conditions is derived for the optimality of a set of parameters and control based on the basic idea of Pontryagin et al. [2]. The following section contains the problem formulation and the main result in the form of a theorem. In Section 3 we establish the validity of the theorem. Section 4 contains conclusion and discussions.

2. PROBLEM FORMULATION AND NECESSARY CONDITIONS

Consider a system governed by the following set of VTIE :

$$x^i(t) = g^i(t,\mu) + \int_{t_o}^{t} h^i(t,\tau)\, f^i(\tau,x(\tau),\, u(\tau),\, \omega)\, d\tau \qquad (I)$$

$$i = 1,2,\ \ldots,\ n$$

Here, $u(t) \in D$, a set of admissible controls, and μ and ω are p and q dimensional constant vectors (of parameters) respectively.

The cost functional to be minimized is,

$$J(t_1) = g^o(t_1,\mu) + \int_{t_o}^{t_1} h^o(t_1,\tau)\, f^o(\tau,x(\tau),u(\tau),\, \omega)\, d\tau \qquad (II)$$

The problem is to find a set $\{\mu^*, \omega^*, u^*(t)\}$ of parameters and control such that $\mu^* \in M \subset R^p$, $\omega^* \in W \subset R^q$; $u^*(t) \in D$ and $J(t_1) = g^o(t_1,\mu^*) + \int_{t_o}^{t_1} h^o(t_1,\tau)\, f^o(\tau,x^*(\tau),u^*(\tau),\, \omega^*)\, d\tau$ is a minimum, where $x^*(t)$ is a solution of equations (I) corresponding to parameters μ^*, ω^*, control u^* and with

$x^*(t_1) \in X_1 \subset R^n$ and $t_1 > t_o$, but otherwise arbitrary.

Defining $x^o(t) = g^o(t,\mu) + \int_{t_o}^{t} h^o(t,\tau)\, f^o(\tau, x(\tau), u(\tau), \omega)\, d\tau$

and writing the augmented system in the matrix vector notation we have,

$$x(t) = g(t,\mu) + \int_{t_o}^{t} H(t,\tau)\, f(\tau, x(\tau), u(\tau), \omega)\, d\tau \qquad \text{(III)}$$

where, $H(t,\tau)$ is a diagonal matrix with the diagonal elements $H_{ij}(t,\tau) = h^i(t,\tau)$, $i = 0, 1, \ldots, n$.

In this notation the problem can be stated as follows :

Find the set $\{\mu^*, \omega^*, x^*(t), u^*(t)\}$

such that $\mu^* \in M$, $\omega^* \in W$, $x^*(t_1) \in R^1 x X_1$, $u^*(t) \in D$ and such that $x^{o^*}(t_1)$ is a minimum, where $t_1 > t_o$ is arbitrary and $x^*(t)$ is a solution to equation - (III) with $\{\mu^*, \omega^*, u^*(t)\}$ as the parameter - control set.

For the formulation of the theorem and its proof the following assumptions are made :

(i) The sets M and W are open subsets of R^p and R^q respectively. X_1 is a smooth manifold of dimension r_1.

(ii) D is the class of functions $u(t)$ with the following properties ;

(a) $u(t)$ is defined and is measurable and bounded in some interval $t_o \leq t \leq t_1$, t_1 depending on u.

(b) $u(t) \in U$ a subset of R^m for any $t_o \leq t \leq t_1$

(c) If $u \in D$ and its interval of definition is $[t_o, t_1]$ then for any $t_o < t' \leq t'' < t_1$ and for any $v \in U$ the function,

$$\bar{u}(t) = \begin{cases} u(t) & \text{if } t_o \leq t \leq t' \text{ ~~or~~ and } t'' < t \leq t_1 \\ v & \text{if } t' < t \leq t'' \end{cases}$$

also belongs to D.

(iii) Let $F_x(t)$ and $F_\omega(t)$ denote the matrices

$$\left[\frac{\partial f}{\partial x}(t, x^*(t), u^*(t), \omega^*)\right] \quad \text{and} \quad \left[\frac{\partial f}{\partial \omega}(t, x^*(t), u^*(t), \omega^*)\right]$$

respectively and assume,

(a) $H(t,s)\, F_x(s)'$ to be measurable in (t,s) for $o \leq s \leq t < \infty$ and

$H(t,s)\, F_x(s)' = 0$ if $s > t$

(b) For each real number $K > 0$, $\exists$ a measurable function m, such that

$||H(t,s)\, F_x(s)'|| \leq m(t,s)$, $t_o \leq s \leq t \leq K$

$\int_{t_o}^{t} m(t,s)\, ds$ is continuous in $[t_o, K]$.

(c) For each compact subinterval I of $[t_o, \infty)$ and each $\bar{t}_o$ in $[t_o, \infty)$,

$$\text{Sup.}\left\{ \int_I ||H(t,s)\, F_x(s)' - H(\bar{t}_o, s)\, F_x(s)'||\, ds \right\} \to o$$

as $t \to \bar{t}_o^{+}$

Let T be an operator defined by

$$z(t) = T\, z(t) = \int_{t_o}^{t} H(t,\tau)\, F_x(\tau)'\, z(\tau)\, d\tau.$$

Then by the above assumptions $[I-T]^{-1}$ exists [16]. Further let, $\psi(t)$ be an absolutely continuous non-zero vector function satisfying the integro-differential equation;

$$\frac{d\ \psi(t)}{dt} = -F_x(t)'\, H(t,t)\, \psi(t) - F_x(t)' \int_{t}^{t_1} \frac{dH(\tau,t)}{d\tau}\, \psi(\tau)\, d\tau$$

Define,

$$\chi(t) := H(t,t)\ \psi(t) + \int_{t}^{t_1} \frac{dH(\tau,t)}{d\tau}\, \psi(\tau)\, d\tau$$

$$K(t) := f(t,x(t),u(t),\omega^*)'\, H(t,t)\ \psi(t) + \int_{t_o}^{t} f(\tau,x(\tau),u(\tau),\omega^*)'\, \frac{dH(t,\tau)}{dt}\, \psi(t)\, d\tau$$

and

$$\mathcal{H}(t,x^*(t),\psi(t),u(t),\omega^*) = f(t,x^*(t),u(t),\omega^*)'\, \chi(t)$$

We state the necessary conditions for the above problem as follows :

<u>Theorem-1</u> :

If $(\mu^*, \omega^*, x^*(t), u^*(t))$ is an optimal solution to the above problem then there exists a non-zero absolutely continuous vector function $\psi(t)$ and the functions χ , K and $\mathcal{H}$ as defined above such that the following conditions are satisfied;

(i) $\mathcal{H}(t,x^*(t),\ \psi(t),\ u^*(t),\ \omega^*)$

$$\leq \mathcal{H}(t,x^*(t),\ \psi(t),v,\ \omega^*)$$

$\forall\ v \in U$ and almost all $t \in [t_o, t_1]$.

(ii) $\psi^o_{t_1}(t_1) \leq o$

(iii) $\int_{t_o}^{t_1} F_\omega(t)' \ \chi(t)\, dt = \underline{o}$

(iv) $K(t_1) = -[\frac{\partial g}{\partial t}(t_1, \mu^*)]' \ \psi(t_1)$

(v) $[\frac{\partial g}{\partial \mu}(t_o, \mu^*)]' \ \psi(t_o) + \int_{t_o}^{t_1} [\frac{d}{dt} \ \frac{\partial g}{\partial \mu}(t, \mu^*)]' \ \psi(t)\, dt = o$

(vi) $\sum_{i=1}^{n} \psi^i(t_1)\ \eta^i = o$, where η is an arbitrary tangent vector to X_1 at $x^*(t_1)$.

3. PROOF OF THEOREM-1 :

For the proof of the theorem we will require the following definition of a regular point of the control function :

Definition :

For any measurable function $v(t)$, $a \leq t \leq b$ with values in a set U a point τ in (a,b) is said to be a regular point of $v(t)$ if for any neighbourhood N of $v(\tau)$ in U and for any sequence of intervals I satisfying $\mathcal{M}(I) \to o$, $\tau \in I$ we have $\mathcal{M}(v^{-1}(N) \cap I)/\mathcal{M}(I) \to 1$, where $\mathcal{M}$ is the Lebesgue measure on $[a,b]$.

Almost all points of any measurable function $v(t)$ are regular. Furthermore if $g(t,v)$ is a continuous function for $a \leq t \leq b$, then for any regular point τ of $v(\tau)$ and for any real number α and β,

$$\int_{\tau+\alpha\varepsilon}^{\tau+\beta\varepsilon} g(t,v(t))\,dt = \varepsilon(\beta-\alpha)\, g(\tau,v(\tau)) + o(\varepsilon)$$

where $o(\varepsilon)$ has its usual meaning.

Let us establish the variational formula,

$$\bar{x}(t) = x^*(t) + \varepsilon \;\Delta x^*(t) + o(\varepsilon)$$

where, $\bar{x}(t)$ is a solution of equation - (III) corresponding to the perturbed parameters, control and terminal time.

The perturbations are given by

$$\bar{\mu} = \mu^* + \varepsilon \;\delta\mu$$

$$\omega = \omega^* + \varepsilon \;\delta\omega$$

$$\bar{u}(t) = \begin{cases} u^*(t) & \text{if} \quad t \notin I_j \,, \;\forall\; j \\ v_j & \text{if} \quad t \in I_j \,, \;\forall\; j \end{cases}$$

$$\bar{t}_1 = t_1 + \varepsilon \;\delta t$$

where, $t_o < \tau_1 < \tau_2 < \cdots < \tau_s < t_1$ are arbitrary regular points of $u^*(t)$ and I_j denotes the interval $\tau_j - \varepsilon\,\delta t_j < t < \tau_j$, $\delta t_j \geq o$ and $\varepsilon > o$ is sufficiently small such that I_j are disjoint intervals.

Let, $z(t) := \bar{x}(t+\varepsilon\;\delta t) - x^*(t)$

Then,

$$z(t) = g(t+\varepsilon\;\delta t,\; \mu^* + \varepsilon\,\delta\mu) - g(t,\mu^*)$$
$$+ \int_{t_o}^{t+\varepsilon\;\delta t} H(t+\varepsilon\,\delta t,\tau)\; f(\tau,\bar{x}(\tau),\bar{u}(\tau),\;\bar{\omega})\; d\tau$$
$$- \int_{t_o}^{t} H(t,\tau)\; f(\tau,x^*(\tau),u^*(\tau),\;\omega^*)\; d\tau$$

or

$$z(t) = [\frac{\partial g(t,\mu^*)}{\partial \mu}] \varepsilon \delta\mu + [\frac{\partial g}{\partial t}(t,\mu^*)] \varepsilon \delta t +$$

$$\sum_{\tau_k < t} H(t,\tau_k) [f(\tau_k,x^*(\tau_k), v(\tau_k), \omega^*) -$$

$$f(\tau_k,x^*(\tau_k), u^*(\tau_k), \omega^*)] \varepsilon \delta t_k +$$

$$H(t,t) f(t,x^*(t),u^*(t), \omega^*) \varepsilon \delta t + \int_{t_o}^{t} H(t,\tau) F_\omega(\tau) d\tau \varepsilon \delta\omega$$

$$+ \int_{t_o}^{t} \frac{dH}{dt}(t,\tau) f(\tau,x^*(\tau),u^*(\tau), \omega^*) d\tau \varepsilon \delta t +$$

$$\int_{t_o}^{t} H(t,\tau) F_x(\tau)' z(\tau) d\tau + o(\varepsilon)$$

Let,

$$G(t) = [\frac{\partial g}{\partial \mu}(t,\mu^*)] \delta\mu + [\frac{\partial g}{\partial t}(t,\mu^*)] \delta t$$

$$+ \sum_{\tau_k < t} H(t,\tau_k) [f(\tau_k,x^*(\tau_k),v(\tau_k), \omega^*)$$

$$- f(\tau_k,x^*(\tau_k), u^*(\tau_k), \omega^*)] \delta t_k$$

$$+ H(t,t) f(t,x^*(t),u^*(t), \omega^*) \delta t +$$

$$\int_{t_o}^{t} H(t,\tau) [\frac{\partial f}{\partial \omega}(\tau,x^*(\tau),u^*(\tau), \omega^*)] d\tau \delta\omega$$

$$+ \int_{t_o}^{t} \frac{dH}{dt}(t,\tau) f(\tau,x^*(\tau),u^*(\tau), \omega^*) d\tau \delta t$$

Witht this, z(t) can be written as

$$z(t) = T\, z(t) + \varepsilon\, G(t) + o(\varepsilon)$$

or

$$(I-T)\, z(t) = \varepsilon\, G(t) + o(\varepsilon)$$

or
$$z(t) = \varepsilon\, G(t) + \sum_{k=1}^{\infty} T^k\, \varepsilon\, G(t) + o(\varepsilon)$$
$$= \varepsilon\, [G(t) + \sum_{k=1}^{\infty} T^k\, G(t)] + o(\varepsilon)$$

From this,

$$\Delta x^*(t) = G(t) + \sum_{k=1}^{\infty} T^k\, G(t)$$

It is seen that $\Delta x^*(\tau')$ depends on $a: = \{\delta\mu, \delta\omega, \tau_i, v_i, \delta t_i, \delta t, \tau'\}$. Fixing τ' and using the same points τ_i and v_i we get $\Delta x^*_a = \lambda' \Delta x^*_{a'} + \lambda'' \Delta x^*_{a''}$, wherever $a = \lambda' a' + \lambda'' a''$. If we take all possible a, the vector $\Delta x^* = \Delta x^*_a$ will fill out a cone $K_{\tau'}$ in the space $X^*_{\tau'}$ ($X^*_{\tau'}$ is the space R^{n+1} translated to $x^*(\tau')$). It can be proved that $K_{\tau'}$ is a convex cone in $X^*_{\tau'}$ and is denoted by the cone of attainability [2].

$K_{\tau'}$ does not fill out the entire space R^{n+1}, as the ray $L_{\tau'}$, starting at $x(\tau')$ and going to the negative of x^o - axis does not belong to the interior of $K_{\tau'}$, as this will violate the optimality condition. Therefore there exists a hyperplane of support Γ to $K_{\tau'}$ at its vertex. Thus there exists a vector l such that for any vector Δx^* the inequalities,

$$(l, \Delta x^*) \leq o \quad , \quad \Delta x^* \in K_{\tau'}$$

$(l,L) \geq 0$, for some $L \notin K_{\tau}$, are satisfied. Let us take $\psi(t_1) = 1$. Indeed we can take L as $(-1- |g^{o}(t_o,\mu^*)| ,0,\ldots 0)$. This gives,

$$\psi^{o}(t_1) \leq 0.$$

This proves condition (ii).

We now prove condition (i).

Let us choose $s = 1$, $\delta t_1 = 1$, $\delta\omega = o$, $\delta\mu = o$ and $\delta t = o$.

Define,

$$B(t) := \int_{t_o}^{t} H(t,\tau)\, F_x(\tau)\, B(\tau)\, d\tau + G(t)$$

With, $B(t_o) = 0$.

This implies

$$G(t) = 0 \quad \text{for} \quad t_o \leq t < \tau_1$$

and $$B(t) = 0 \quad \text{for} \quad t_o \leq t < \tau_1$$

For $t > \tau_1$

$$G(t) = H(t,\tau)C \quad \text{where,}$$

$$C = f(\tau_1,x^*(\tau_1),\, v_1,\, \omega^*) - f(\tau_1,x^*(\tau_1),u^*(\tau_1),\, \omega^*)$$

It follows that

$$B(t) = \int_{\tau_1}^{t} H(t,\tau)\, F_x(\tau)\, B(\tau)\, d\tau + H(t,\tau_1)\, C \qquad \tau_1 < t \leq t_1$$

-ferentiating,

$$\frac{dB(t)}{dt} = H(t,t)\ F_x(t)\ B(t) + \int_{\tau_1}^{t} \frac{dH(t,\tau)}{dt} F_x(\tau)\ B(\tau)\ d\tau$$
$$+ \frac{dH}{dt}(t,\tau_1)C$$

We have,

$$\int_{\tau_1}^{t_1} \frac{d}{dt} < B(t),\ \psi(t) > dt = \int_{\tau_1}^{t_1} [< \frac{dB(t)}{dt}, \psi(t) >$$
$$+ < B(t),\ \frac{d\psi(t)}{dt} >]\ dt$$

$$= \int_{\tau_1}^{t_1} [< H(t,t)\ F_x(t)\ B(t),\ \psi(t) > +$$
$$< \int_{\tau_1}^{t} \frac{dH(t,\tau)}{dt} F_x(\tau)\ B(\tau)\ d\tau,\ \psi(t) > +$$
$$< \frac{dH}{dt}(t,\tau_1)C,\ \psi(t) >]\ dt -$$
$$\int_{\tau_1}^{t_1} [< B(t),\ F_x(t)'\ H(t,t)\ \psi(t) > +$$
$$< B(t),\ F_x(t)' \int_{t}^{t_1} \frac{dH(\tau,t)}{d\tau}\ \psi(\tau)\ d\tau >]\ dt.$$
$$= \int_{\tau_1}^{t_1} [< \int_{\tau_1}^{t} \frac{dH(t,\tau)}{dt} F_x(\tau)\ B(\tau)\ d\tau, \psi(t) > -$$
$$< B(t),\ F_x(t)' \int_{t}^{t_1} \frac{dH(\tau,t)}{d\tau}\ \psi(\tau)\ d\tau >]\ dt$$
$$+ \int_{\tau_1}^{t_1} < \frac{dH}{dt}(t,\tau_1)C,\ \psi(t) > dt$$

$$= \int_{\tau_1}^{t_1} \int_{\tau}^{t_1} < \frac{dH}{dt}(t,\tau)\ F_x(\tau)\ B(\tau),\ \psi(t) > dt\ d\tau$$

$$- \int_{\tau_1}^{t_1} < B(t),\ F_x(t)' \int_{t}^{t_1} \frac{dH}{d\tau}(\tau,t)\ \psi(\tau)\ d\tau > dt$$

$$+ \int_{\tau_1}^{t_1} < \frac{dH}{dt}(t,\tau_1)C,\ \psi(t) > dt$$

$$= \int_{\tau_1}^{t_1} < \frac{dH}{dt}(t,\tau_1)C,\ \psi(t) > dt$$

Thus,

$$\int_{\tau_1}^{t_1} \frac{d}{dt} < B(t),\ \psi(t) > dt = \int_{\tau_1}^{t_1} < \frac{dH}{dt}(t,\tau_1)C,\ \psi(t) > dt$$

The above implies that

$$< B(t_1),\ \psi(t_1) > - < B(\tau_1+o),\ \psi(\tau_1) >$$

$$= < \Delta x^*,\ \psi(t_1)> - < H(\tau_1,\tau_1)C,\ \psi(\tau_1)>$$

$$= \int_{\tau_1}^{t_1} < \frac{dH}{dt}(t,\tau_1)C,\ \psi(t) > dt$$

or

$$< H(\tau_1,\tau_1)C,\ \psi(\tau_1) > + \int_{\tau_1}^{t_1} < \frac{dH}{dt}(t,\tau_1)C,\ \psi(t) > dt$$

$$= < \Delta x^*,\ \psi(t_1) > \leq 0$$

Substituting for C in the above inequality and from the definitions of χ and $\mathfrak{H}$ we have, since τ_1 is any regular point of $u^*(t)$ and $v_1 \in U$ is arbitrary the condition (i) is established.

For the verification of condition (iii) we choose the following set of variations :

$$\bar{u}(t) = u^*(t)$$

$$\delta\mu = o$$

$$\delta t = o$$

and $\delta\omega$ arbitrary.

For this set of variations we have,

$$G(t) = \int_{t_o}^{t} H(t,\tau)\, F_\omega(\tau)\, \delta\omega \; d\tau$$

Defining B(t) as before and substituting for G from the above equations,

$$B(t) = \int_{t_o}^{t} H(t,\tau)\, [F_x(\tau)\, B(\tau) + F_\omega(\tau)\, \delta\omega]\, d\tau$$

$$\begin{aligned}\frac{dB(t)}{dt} &= H(t,t)\, [F_x(t)\, B(t) + F_\omega(t)\, \delta\omega] \\ &\quad + \int_{t_o}^{t} \frac{dH(t,\tau)}{dt}\, [F_x(\tau)\, B(\tau) + F_\omega(\tau)\, \delta\omega]\, d\tau \\ &= H(t,t)\, F_x(t)\, B(t) + \int_{t_o}^{t} \frac{dH(t,\tau)}{dt}\, F_x(\tau)\, B(\tau)\, d\tau \\ &\quad + H(t,t)\, F_\omega(t)\, \delta\omega + \int_{t_o}^{t} \frac{dH(t,\tau)}{dt}\, F_\omega(\tau)\, \delta\omega \; d\tau\end{aligned}$$

$$\begin{aligned}\int_{t_o}^{t_1} \frac{d}{dt} < B(t), \psi(t) > dt &= < \Delta x^*, \psi(t_1)> - < B(t_o+o), \psi(t_o) > \\ &= < \Delta x^*, \psi(t_1) > \leq 0\end{aligned}$$

This implies

$$\int_{t_o}^{t_1} < \frac{dB}{dt}(t), \psi(t) > dt + \int_{t_o}^{t_1} < B(t), \frac{d\psi}{dt}(t) > dt \leq 0$$

$$\Longrightarrow \int_{t_o}^{t_1} [< H(t,t)\, F_\omega(t)\, \delta\omega, \psi(t) >$$

$$+ < \int_{t_o}^{t} \frac{dH}{dt}(t,\tau)\, F_\omega(\tau)\, \delta\omega\, d\tau, \psi(t) >]\, dt \leq 0$$

or

$$\int_{t_o}^{t_1} < H(t,t)\, F_\omega(t)\, \delta\omega , \psi(t) > dt +$$

$$\int_{t_o}^{t_1} \int_{t_o}^{t} < \frac{dH}{dt}(t,\tau)\, F_\omega(\tau)\, \delta\omega\, d\tau, \psi(t) > d\tau\, dt \leq 0$$

or

$$\int_{t_o}^{t_1} < H(t,t)\, F_\omega(t)\, \delta\omega, \psi(t) > dt +$$

$$\int_{t_o}^{t_1} \int_{t}^{t_1} < \frac{dH}{d\tau}(\tau,t)\, F_\omega(t)\, \delta\omega , \psi(\tau) > d\tau\, dt \leq 0$$

or

$$\delta\omega' \int_{t_o}^{t_1} [F_\omega(t)'\, H(t,t)\, \psi(t) + F_\omega(t)' \int_{t}^{t_1} \frac{dH}{d\tau}(\tau,t)\, \psi(\tau)\, d\tau]\, dt \leq 0$$

or

$$\delta\,\omega' \int_{t_0}^{t_1} F_\omega(t)' \,[H(t,t)\,\psi(t) + \int_{t}^{t_1} \frac{dH}{d\tau}(\tau,t)\,\psi(\tau)\,d\tau]\,dt \leq 0$$

or

$$\delta\,\omega' \int_{t_0}^{t_1} F_\omega(t)' \quad \chi(t)\,dt \leq 0$$

Hence, condition (iii) is obviously satisfied since W is an open subset of R^q.

Next to verify condition (iv) we choose the set of variations as;

$$\bar{u}(t) = u^*(t)$$

$$\delta\mu = 0$$

$$\delta\omega = 0$$

and δt arbitrary.

With this set of variations we obtain,

$$z(t) = [\frac{\partial g}{\partial t}(t,\mu^*) + \int_{t_0}^{t} \frac{dH}{dt}(t,\tau)\, f(\tau,x^*(\tau),u^*(\tau),\, \omega^*)d\tau$$

$$+ H(t,t)\, f(t,x^*(t),\, u^*(t),\, \omega^*)]\;\varepsilon\;\delta t + o(\varepsilon)$$

We have,

$$\Delta x^*(t) = \delta t[\frac{\partial g}{\partial t}(t,\mu^*) + \int_{t_0}^{t} \frac{dH}{dt}(t,\tau)\, f(\tau)d\tau + H(t,t)f(t)]$$

where for notational convenience $f(\cdot,x^*(\cdot),\, u^*(\cdot),\, \omega^*)$ is represented as $f(\cdot)$

The above equation implies,

$$\Delta x^*(t_1) = \delta t \; [\; \frac{\partial g}{\partial t}(t_1, \mu^*) + \int_{t_o}^{t_1} \frac{dH}{dt}(t_1, \tau)\; f(\tau)\; d\tau$$

$$+ \; H(t_1, t_1)\; f(t_1)]$$

Thus,

$$< \Delta x^*(t_1), \; \psi(t_1)> \;\; = \delta t \; \Big[[\frac{\partial g}{\partial t}(t_1, \mu^*)]' \;\; \psi(t_1)$$

$$+ \int_{t_o}^{t_1} f(\tau)' \; \frac{dH}{dt}(t_1, \tau) \; \psi(t_1) \; d\tau + f(t_1)' \; H(t_1, t_1) \;\; \psi(t_1) \Big]$$

$$\leq 0$$

or

$$\delta t \; \Big[[\; \frac{\partial g}{\partial t}(t_1, \mu^*)]' \; \psi(t_1) \; + \; K(t_1) \Big] \leq 0$$

This implies

$$K(t) \; = - \; [\; \frac{\partial g}{\partial t}(t_1, \mu^*)]' \;\; \psi(t_1)$$

Condition (v) can be similarly verified by choosing the following set of variations :

$$\bar{u}(t) \; = u^*(t)$$

$$\delta\omega \;\; = 0$$

$$\delta t \;\; = 0$$

$\delta\mu$ is arbitrary.

We derive below the last condition.

Let $(\mu^*, \; \omega^*, \; x^*(t), \; u^*(t), \; t_1)$ be the optimal set. Let $x^*(t_1) = \hat{x}_1 = (x^o(t_1), \; x_1)$. Let $\bar{T}_1$ be the set of tangent vectors of the manifold X_1 at the point x_1 and let $\mathcal{J}_1$ be the r_1-dimensional manifold in R^{n+1} consisting of all points of

the form $(x^o(t_1), x)$, where $x \in \bar{T}_1$. Let us pass a ray through each point of $\mathcal{J}_1$ in the direction of negative x^o-axis and denote by Q the point set generated by all these rays. The set Q is a half-space on an (r_1+1)-dimensional space and its boundary points belong to $\mathcal{J}_1$. Since $x^*(t_1)$ is an optimal point; K_{t_1} and relative interior of Q have empty intersection and hence Q and K_{t_1} are separated [43]. This implies that there exist constants $C_o, \ldots, C_n$ such that K_{t_1} is in the half-space $\sum_{\alpha=0}^{n} C_\alpha x^\alpha \leq 0$, where $(x^o, x^1, \ldots, x^n)$ are coordinates in $X^*_{t_1}$ and Q is in the half-space $\sum_{\alpha=0}^{n} C_\alpha x_\alpha \geq 0$. The (r_1+1) - dimensional manifold $\mathcal{J}_1 \subset Q$ lies in the half-space $\sum_{\alpha=o}^{n} C_\alpha x^\alpha \geq 0$ and consequently in the hyper-plane

$$\sum_{\alpha=o}^{n} C_\alpha x^\alpha = o.$$ With this, $\psi(t_1)$ can be taken as equal to

$(C_o, C_1, \ldots, C_n)$. Since $(o, \eta_1, \ldots, \eta_n)$ is an element of $\mathcal{J}_1$ for any tangent vector $(\eta_1, \ldots, \eta_n)$ to X_1, the relation

$$\sum_{\alpha=o}^{n} \psi_\alpha(t_1)\, x^\alpha = o \text{ reduces to } \sum_{i=1}^{n} \psi_i(t_1)\, \eta^i = o.$$

Hence the validity of the theorem is established.

4. CONCLUSION

In the above theorem the sets M and W were assumed to be open subsets of R^p and R^q respectively. In case the sets M and W are smooth closed subsets, a similar theorem can be established, where the equalities (iii) and (v) in the

theorem 1 are replaced by the following inequalities respectively.

$$\text{(iii)} \quad \delta\omega' \int_{t_o}^{t_1} F_\omega(t)' \; \chi(t)\, dt \leq 0$$

$$\text{(v)} \quad \delta\mu' \left[\frac{\partial g}{\partial \mu}(t_o, \mu^*) \right]' \psi(t_o) +$$

$$\delta\mu' \int_{t_o}^{t_1} \left[\frac{d}{dt} \frac{\partial g}{\partial \mu}(t, \mu^*) \right]' \psi(t)\; dt \leq 0$$

Here the variations $\delta\mu$ and $\delta\omega$ are restricted to be contained in the tangent cones of $(M-\mu^*)$ and $(W-\omega^*)$, at the origin, respectively.

CHAPTER IV

OPTIMAL CONTROL OF SYSTEMS DESCRIBED BY DELAY-DIFFERENTIAL EQUATIONS

1. INTRODUCTION

Application of Pontryagin's maximum principle to the optimization of time-delay systems results in a system of coupled two-point-boundary value problems with both delay and advanced terms [2]. Appropriate numerical techniques have to be used for solving these problems.

Eller et al. [31] have presented a technique for solving the optimization problem with quadratic cost functional that requires simultaneous solution of $2n^2+n$ equations for an n^{th} order system using a nontrival numerical technique [32]. This method has also the disadvantage that it generates open-loop control. Another approach for the optimization of linear time-delay system is to transform the problem in to an equivalent linear non-delay problem. This approach, as developed by Slater and Wells [33], allows only the integral of a quadratic form of the control, but not the state in the cost functional.

The computation of sub-optimal controls, for linear time-delay systems has been considered by several authors. A section of these results is based on a continuation or sensitivity analysis concept. Inoue et al. [34] have used this approach to

obtain a sub-optimal control for stationary linear systems with small delay in the state. They have expanded the control in a MacLaurin series in the delay and obtained the series coefficients from the solution of simple TPBVPS. Jamshidi and Malek-Zavarei [35] used a MacLaurin series expansion of the control in a sensitivity parameter and obtained the series coefficients from non-delay system computations. The sub-optimal control they obtained is in the form of an exact feedback term and truncated series forward term for stationary linear system with time-delay. The result has been extended to the case of multiple delay in state and control by Malek-Zavarei [36].

Gracovetsky and Vidyasagar [38] have proposed an iterative scheme for the optimization of linear non-stationary system with delay in state. They have extended the result to optimize systems described by nonlinear delay differential equations, with delay appearing in state [40] and also to neutral systems [39]. Recently Malek-Zavarei [37] has extended the results to optimize a more general class of delay systems. He has considered linear system with multiple delay in state and control. He also extends the result to nonlinear delay systems with time-varying delay in state and control.

Each iteration in Gracovetsky et al.'s scheme consists of solving an optimization problem with ordinary differential equations as constraints. In the k^{th} optimization problem the delay term in the system equation is substituted by the solutions

obtained from the $(k-1)^{th}$ iteration. In each iteration one requires to solve 2n - ordinary differential equations with n-initial and n-final conditions given, for a n^{th}-order system. This indeed is an easy task as compared to other methods. The method also results in a control which is partially closed-loop and partially open-loop.

For the convergence proof of the sequence of optimization problems, Gracovetsky et al. have proposed a lemma based on the Picard's method for existence of solution of ordinary differential equations. The lemma seems to be inadequate to ascertain any conclusion regarding the convergence of the sequence of optimization problems. Each term of the sequence consists of equations with both initial and final values given, thereby forming a TPBVP and in the lemma each term of the sequence comprises of an initial value problem.

Further as it is clearly seen, if at all the sequence so obtained converges, in the limit it does not, in general, satisfy the maximum principle corresponding to the delay system. One is totally uncertain, whether the control obtained through the proposed optimization sequence is sub-optimal or not. The method cannot be applied with any confidence, though in some particular cases it may give near optimal controls.

In this chapter we present an iterative scheme to determine the optimal control for a system described by linear time-delay equations with quadratic cost functional. The control obtained

is partly closed-loop and partly open-loop. The convergence proof for the iterative scheme is given for systems with certain restrictions. The iteration scheme is expected to converge, in general, wherever the TPBVP corresponding to the system has a solution.

The following two sections contain the main results of the chapter. In Section 2 we prove a lemma essential for the proof of the theorem in Section 3. Section 4 contains a numerical example illustrating the method. The last section (Section 5) contains extensions and discussions.

2. Lemma :

Let,

$$\dot{g}(t) = -A(t)'\, g(t) - h(t, x(t-\delta), x(t+\delta), g(t+\delta))$$

$$\dot{x}(t) = A(t)\, x(t) + f(t, x(t-\delta), g(t)) \tag{2.1}$$

$x \in R^n$, $g \in R^n$ and $\forall t \in [0,T]$

with boundary conditions,

$$x(t) = \eta(t) \qquad -\delta \leq t \leq 0$$

$$x(t) = 0 \qquad T < t \leq T + \delta$$

$$g(t) = 0 \qquad T \leq t \leq T + \delta$$

be a given system; f and h are Lipschitzian and A is a piecewise continuous matrix of appropriate dimension. η is an arbitrary given function in $C_n[-\delta, 0]$; the space of all continuous functions in $[-\delta, 0]$ with range in R^n.

Then the sequence of vector functions $\{x^i, g^i\}_{i\in N}$ defined by,

$$g^o(t) = 0 \qquad 0 \leq t \leq T + \delta$$

$$x^o(t) = \Phi(t,o)\ \eta(o) \qquad 0 \leq t \leq T$$

$$x^o(t) = \eta(t) \qquad -\delta \leq t \leq 0$$

$$x^o(t) = 0 \qquad T < t \leq T + \delta$$

$$g^i(t) = \int_t^T [\Phi(\tau,T)]^{-1}\ h(\tau,x^{i-1}(\tau-\delta),x^{i-1}(\tau+\delta),g^{i-1}(\tau+\delta))\, d\tau \qquad \forall t \in [0,T]$$

$$x^i(t) = \Phi(t,o)\ \eta(o) + \int_0^t \Phi(t,\tau)\ f(\tau,x^{i-1}(\tau-\delta),g^i(\tau))\, d\tau \qquad \forall t\in[0,T]$$

$$g^i(t) = 0\ ; \quad T \leq t \leq T + \delta$$

$$x^i(t) = 0\ ; \quad T < t \leq T + \delta$$

$$x^i(t) = \eta(t);\ -\delta \leq t \leq 0$$

(where $\Phi(t,\tau)$ is the state transition matrix corresponding to $A(t)$) converges uniformly to the solution of equations (2.1), if the Lipschitz constants for the functions h and f are sufficiently small.

Proof :

Let,

$$M: = \sup_{t,\tau\in[0,T]} \|\Phi(t,\tau)\|$$

$$L: = \sup_{t\in[-\delta,0]} \|\eta(t)\|$$

$$N: = \mathrm{Max}[F,H]$$

where, F and H are the Lipschitz constants of the functions f and h respectively, i.e., for every (t,x_1,y_1,z_1) and (t,x_2,y_2,z_2) in the domain of h,

$$\|h(t,x_1,y_1,z_1) - h(t,x_2,y_2,z_2)\| \leq H(\|x_1-x_2\| + \|y_1-y_2\| + \|z_1-z_2\|)$$

and for every (t,x_1,y_1) and (t,x_2,y_2) in the domain of f,

$$\|f(t,x_1,y_1) - f(t,x_2,y_2)\| \leq F.(\|x_1-x_2\| + \|y_1-y_2\|).$$

Here, $\|\cdot\|$ denotes the Euclidean norm.

Further let, $\alpha: = MNT \leq 1/6$

Now,

$$\|g^1(t) - g^0(t)\| \leq MN \int_t^T (\|x^0(\tau-\delta)\| + \|x^0(\tau+\delta)\|)\, d\tau$$

$$\leq 2\ MNL \int_t^T d\tau \leq 2L\alpha$$

$$\|x^1(t) - x^0(t)\| \leq MN \int_0^t (\|x^0(\tau-\delta)\|+\|g^1(\tau)\|)\, d\tau$$

$$\leq LMNt + 2L\alpha^2 \leq 3L\alpha$$

$$\| g^2(t) - g^1(t) \| \leqslant MN \int_t^T (\| x^1(\tau-\delta) - x^0(\tau-\delta) \|$$

$$+ \| x^1(\tau+\delta) - x^0(\tau+\delta) \| + \| g^1(\tau+\delta) - g^0(\tau+\delta) \|) \, d\tau$$

$$\leqslant 3\ MNT \cdot 3L\alpha \leqslant L \cdot (3\alpha)^2$$

$$\| x^2(t) - x^1(t) \| \leqslant MN \int_0^t (\| x^1(\tau-\delta) - x^0(\tau-\delta) \| + \| g^2(\tau)$$

$$- g^1(\tau) \|) \, d\tau \leqslant MNt(3L\alpha + L(3\alpha)^2)$$

$$\leqslant L(3\alpha)^2$$

Similarly by induction one can show,

$$\| g^k(t) - g^{k-1}(t) \| \leqslant L.(3\alpha)^k \leqslant L \cdot \frac{1}{2^k}$$

$$\| x^k(t) - x^{k-1}(t) \| \leqslant L.(3\alpha)^k \leqslant L. \frac{1}{2^k}$$

Applying the triangle inequality, for any r

$$\| g^{k+r}(t) - g^k(t) \| \leqslant L. \frac{1}{2^{k+1}} \left(1 + \tfrac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{r-1}}\right)$$

$$\leqslant L. \frac{1}{2^k}$$

$$\| x^{k+r}(t) - x^k(t) \| \leqslant L. \frac{1}{2^k}$$

Therefore the sequence is uniformly convergent. ✓

3. PROBLEM STATEMENT AND METHOD OF SOLUTION

Consider the linear delay system,

$$\dot{x}(t) = A(t)\ x(t) + C(t)\ x(t-\delta) + B(t)\ u(t), \forall t \geqslant 0 \tag{3.1}$$

$$x(t) = \eta(t), \quad -\delta \leq t \leq o$$

$x \in R^n$, $u \in R^m$ and the system matrices are piecewise continuous with appropriate dimensions. $\eta \in C_n[-\delta, o]$ is an arbitrary given function,

Let,

$$J := \tfrac{1}{2}\, x(T)'\, F\, x(T) + \tfrac{1}{2} \int_0^T [x(t)'\, Q(t)\, x(t) + u(t)'\, R(t)\, u(t)]\, dt$$

where,

F, Q and R are matrices of appropriate dimensions. F is positive semi-definite, Q is symmetric positive semi-definite and piecewise continuous and R is symmetric positive definite and piecewise continuous.

The problem is to find a control $u(t)$, $o \leq t \leq T$, which for fixed final time T and free final state $x(T)$ minimizes the functional J.

<u>Theorem</u> :

Consider the sequence,

$$g^o(t) = o \qquad o \leq t \leq T + \delta$$

$$\dot{x}^o(t) = [A(t) - S(t)\, K(t)]x^o(t), \; x^o(o) = \eta(o), \; t \in [0,T]$$

$$x^o(t) = \eta(t), \quad -\delta \leq t \leq o$$

$$\dot{g}^i(t) = -[A(t) - S(t)\, K(t)]'\, g^i(t) - K(t)\, C(t)\, x^{i-1}(t-\delta)$$
$$- C\,(t+\delta)'\, K(t+\delta)\, x^{i-1}(t+\delta) - C(t+\delta)'\, g^{i-1}(t+\delta)$$
$$o \leq t \leq T-\delta$$

$$\dot{g}^i(t) = -[A(t) - S(t)\,K(t)]'\,g^i(t) - K(t)\,C(t)\,x^{i-1}(t-\delta)$$

$$g^i(T) = 0 \qquad\qquad T-\delta \leq t \leq T$$

$$\dot{x}^i(t) = [A(t) - S(t)\,K(t)]\,x^i(t) - S(t)\,g^i(t) + C(t)\,x^{i-1}(t-\delta),$$

$$x^i(t) = \eta(t),\ -\delta \leq t \leq 0. \qquad\qquad 0 \leq t \leq T$$

$i = 1,2, \ldots,$

Here, $S(t) := B(t)\,R^{-1}(t)\,B(t)'$ and $K(t)$ is the symmetric positive definite solution of the matrix Riccati Equation,

$$\dot{K}(t) + K(t)\,A(t) + A(t)'\,K(t) - K(t)\,S(t)\,K(t) + Q(t) = 0 \qquad \forall\, t \in [0,T]$$

with boundary conditions, $K(T) = F$

Let $\Phi(t,\tau)$ be the state transition matrix corresponding to the system matrix $[A(t) - S(t)\,K(t)]$. Then if $||\Phi(t,\tau)||$, $||S(t)||$, $||C(t)||$ and $||K(t)||$ are sufficiently small the sequence $\{x^i\}_{i\in N}$ (where $x^i \in C_n[0,T]$ is a solution of the equations in the i^{th} iteration) converges to the optimal solution x^*, of the problem and the optimal control u^* is obtained as,

$$u^*(t) = -R^{-1}(t)\,B(t)'\,K(t)\,x^*(t) + g(t), \quad 0 \leq t \leq T$$

Here g is the limit of the sequence $\{g^i\}$.

<u>Proof</u> :

The convergence is obvious from the lemma stated in the previous section. Also the limit of the sequence $\{x^i\}$, $\{g^i\}$

satisfy the set of equations,

$$\dot{g}(t) = - [A(t) - S(t) K(t)]' g(t) - K(t) C(t) x^*(t-\delta)$$

$$- C(t+\delta)' K(t+\delta) x^*(t+\delta) - C(t+\delta)' g(t+\delta)$$

$$o \leq t \leq T - \delta$$

$$\dot{g}(t) = - [A(t) - S(t) K(t)]' g(t) - K(t) C(t) x^*(t-\delta)$$

$$T - \delta \leq t \leq T$$

$$g(T) = 0$$

$$\dot{x}^*(t) = [A(t) - S(t) K(t)] x^*(t) - S(t) g(t) + C(t) x^*(t-\delta)$$

$$o \leq t \leq T$$

$$x^*(t) = \eta(t), \quad -\delta \leq t \leq o.$$

Applying the maximum principle to the given system and substituting $[K(t) x^*(t) + g(t)]$ for the auxilary variables along with the matrix Riccati equation one obtains the above set of equation as the sufficient optimality criteria.

4. NUMERICAL EXAMPLE

Consider the example of a typical control problem arising in the chemical and petrolium industries. The block diagram of figure (IV.1) depicts a refining plant.

F_A and F_B represent the feed rate, in pounds per hour, of the raw materials A and B respectively. The reaction in the reactor is exothermic and the reactants cooled to a desired temperature in the heat exchanger. The undesirable products settle out of the reactant mixture in the decanter. Finally

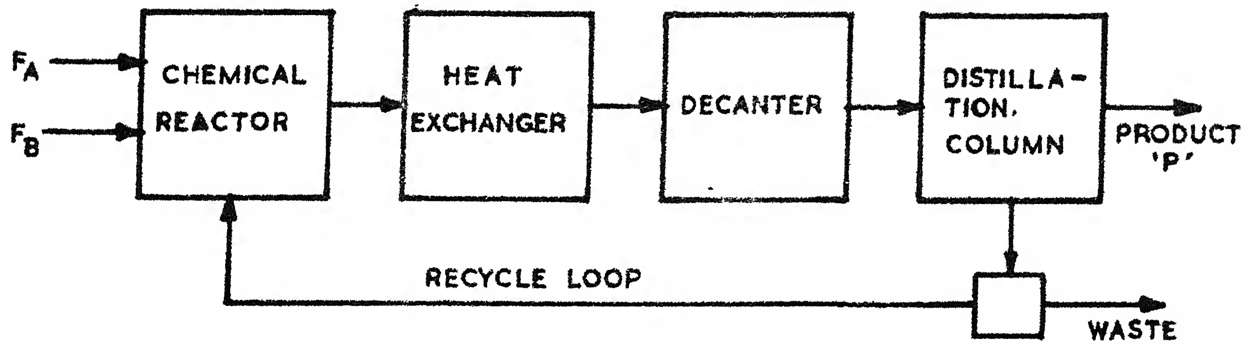

FIG.IV.1 BLOCK DIAGRAM OF THE REFINING PLANT

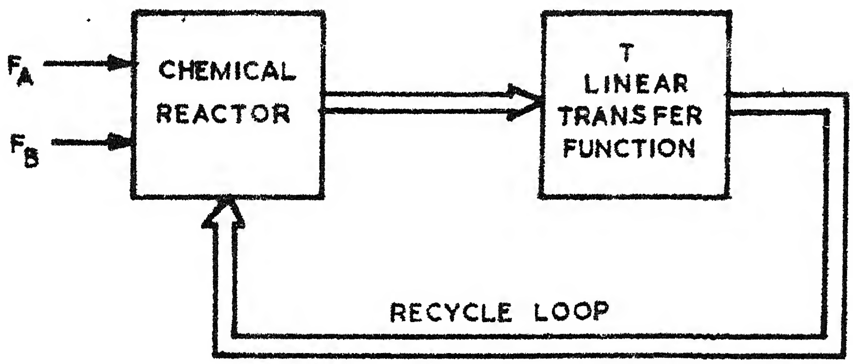

FIG.IV.2 SIMPIFIED BLOCK DIAGRAM OF THE REFINING PLANT

transformation. The block diagram then reduces to one as shown in figure (IV.2).

With numerical values given in [42] and 10 mts. recycle time the linearized model is obtained as [41],

$$\dot{x}(t) = A\ x(t) + C\ x(t-1) + B\ u(t)$$

where the system matrices are,

$$A = \begin{bmatrix} -4.93 & -1.01 & 0.0 & 0.0 \\ -3.20 & -5.30 & -12.8 & 0.0 \\ 6.40 & 0.347 & -32.5 & -1.04 \\ 0.0 & 0.833 & 11.0 & -3.96 \end{bmatrix}$$

$$C = \begin{bmatrix} 1.92 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.92 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.87 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.724 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Here the time unit is 10 mts. The various state and control variables represent the following physical quantities :

x_1 : deviation in the weight composition (dimensionless) of the reactant A from its nominal value.

x_2 : deviation in the weight composition of the reactant B from its nominal value.

x_3 : deviation in the weight composition of an intermediate product.

x_4 : deviation in the weight composition of the product P from its nominal value.

$u_1 := \frac{\delta F_A}{6 V_R}$; $u_2 := \frac{\delta F_B}{6 V_R}$, where V_R is the pound volume of the chemical reactor and δF_A and δF_B are respectively deviation in the feed rate of A and B from their nominal values.

A typical time response of the uncontrolled system, with $\eta_1(t) = 0.1$, $\eta_2(t) = \eta_3(t) = \eta_4(t) = 0$; $\forall\, t \in [-1,0]$, is shown in figure (IV.3). The response is sluggis.

As a large volume of chemicals are being processed and the quality of the product is to be maintained, any deviation from the nominal values can be very costly. This makes the natural response of the system unacceptable and a controller is to be designed to rapidly diminish the deviations in the reactant compositions.

The control functions were computed using the iterative scheme presented in the previous section for the following choice of Q,R,F and T :

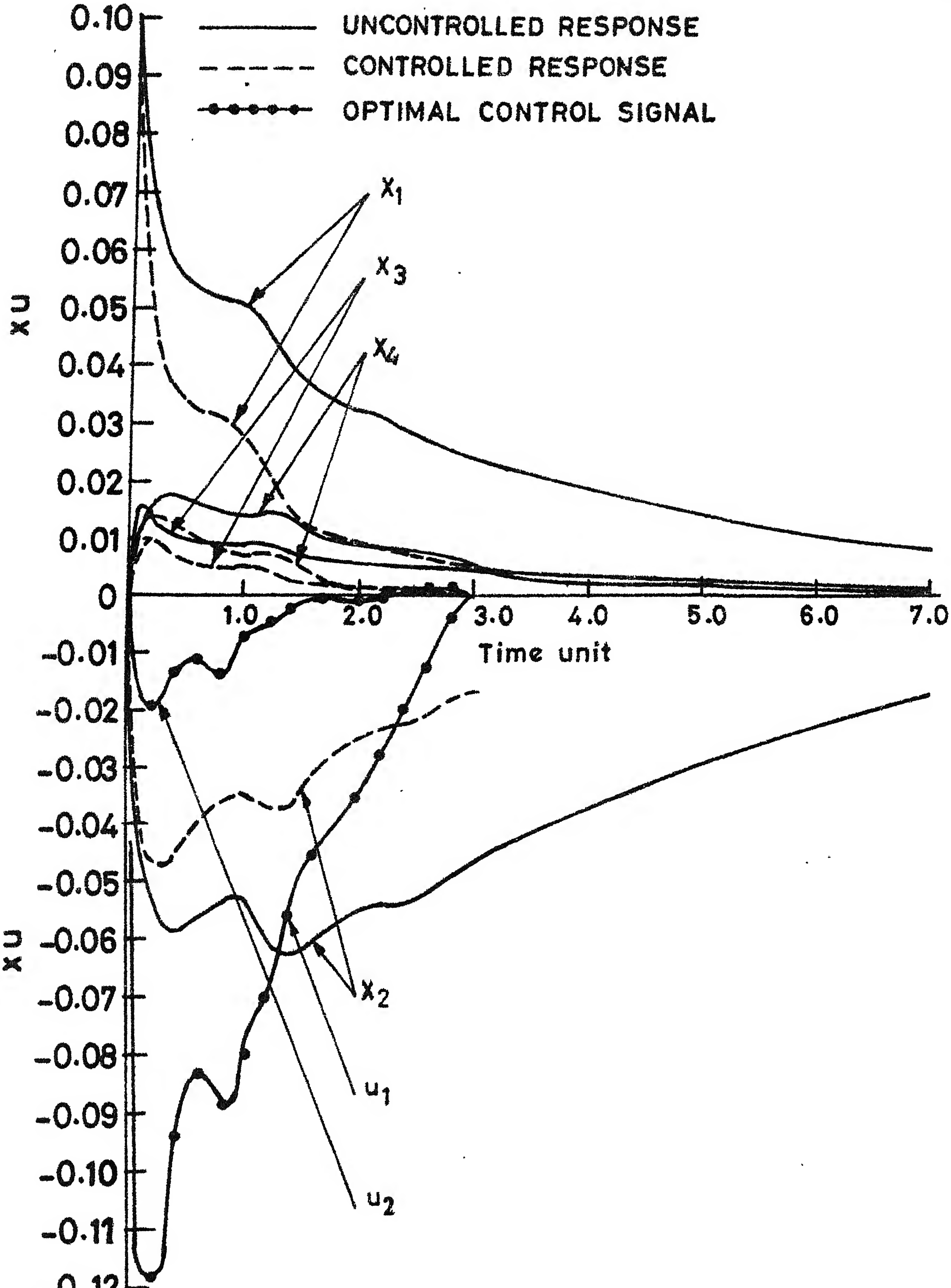

FIG.IV.3 SYSTEM RESPONSE AND OPTIMAL CONTROL SIGNAL

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 100 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad F = [0], \; T = 3 \text{ units}$$

The system response with the control inputs as presented in **Fig.IV.3** depicts the improvement in the system performance.

5. CONCLUSION

The technique developed in this chapter can directly be extended to multiple nonlinear time-varying delays in state variables.

The iterative scheme presented is quite easy to apply as it requires solution of a set of ordinary differential equations with final value and a set of ordinary differential equations with initial value in each iteration. The Riccati equation is required to be solved once only and each iteration step requires the solution of 2n-ordinary differential equations. The example considered depicts that the proposed algorithm works even when the smallness conditions used in the convergence proof is far from being satisfied.

The algorithm requires the storage of the 2n-functions x and g in addition to the n^2-functions K. One can follow the following alternative iteration scheme for the solution of the optimization problem ;

$$\dot{g}^i(t) = -[A(t) - S(t)K(t)]'\, g^i(t) - K(t)\, C(t)\, x^{i-1}(t-\delta), \qquad T-\delta \leq t \leq T$$

$$g^i(T) = 0$$

$$\dot{g}^i(t) = -[A(t) - S(t)K(t)]'\, g^i(t) - K(t)\, C(t)\, x^{i-1}(t-\delta) - C(t+\delta)'\, K(t+\delta) x^{i-1}(t+\delta) - C(t+\delta)'\, g^i(t+\delta); \qquad 0 \leq t \leq T-\delta$$

$$\dot{x}^i(t) = [A(t) - S(t)K(t)]\, x^i(t) - S(t)\, g^i(t) + C(t) x^i(t-\delta), \qquad 0 \leq t \leq T$$

$$x^i(t) = \eta(t), \quad -\delta \leq t \leq 0$$

$i = 1,2,3, \ldots$

This scheme requires the storage of n-functions x and n^2-function K. Further as the renewed value of the advanced term in g is used in solving the equations for g and the renewed value of the delay term in x for solving the equations for x is being used the rate of convergence may be better.

CHAPTER V

DETERMINATION OF OPTIMAL INITIAL FUNCTION AND PARAMETERS IN A SYSTEM DESCRIBED BY DELAY-DIFFERENTIAL EQUATIONS

1. INTRODUCTION

Many physical systems are represented by a system of Delay-Differential Equations (DDE) with adequate accuracy. In such systems various problems (e.g., modelling, model reduction, design problems etc.) may require the determination of certain system parameters as well as the initial functions for a given system of DDE so as to obtain certain desired system behaviour.

In this chapter we consider the problem of determining the initial functions and parameters in a given system of DDE so as to minimize a functional of various system variables. Mathematically the problem considered can be stated as follows :

Problem (P) :

Let,

$$\dot{x}(t) = f(t,x(t)\ ,\ x(t-1),\alpha), \quad t \geq 0 \tag{1.1a}$$

$$x(t) = \eta(t), \quad -1 \leq t \leq 0 \tag{1.1b}$$

$$x \in C_n[-1,T],\ \alpha \in R^m,\ \eta \in C_n[-1,0]$$

be a given system. Here, $C_r[a,b]$ denotes the space of all continuous functions on the interval $[a,b]$ with range in R^r.

Let,

$$J: C_n[-1,0] \times C_n[0,T] \times R^m \to R$$

be a given functional.

The problem is to determine $\eta^* \in C_n[-1,0]$, $x^* \in C_n[0,T]$ and $\alpha^* \in R^m$ such that J attains a minimum among all (η, x, α) satisfying the system equations (1.1) and $\eta \in X \subset C_n[-1,0]$; where X is a l-dimensional subspace of $C_n[-1,0]$ defined as;

$X := \{\eta \in C_n[-1,0] : \eta(t) = \mu . e(t) \;\forall\, t \in [-1,0],\; \mu \in R^{nxl}$ is any nxl constant matrix, $e \in C_l[-1,0]$ and $\{e_i : i = 1,2, \ldots, l\}$ is a given set of linearly independent elements of $C_1[-1,0]$. Here, e_i denotes the i^{th} component of $e\}$

In the above problem we seek the optimal initial function in a finite dimensional subspace of $C_n[-1,0]$. For all practical purposes, the space of continuous functions can be approximated by a finite dimensional subspace of it. If the dimension l of X in problem (P) is chosen sufficiently large we obtain a near-optimal solution for the initial function in the space $C_n[-1,0]$.

This formulation of the problem is motivated by the fact that it reduces the problem of determination of the optimal initial function to that of determining a finite set of parameters.

In Section 2 we discuss the existence of an optimal solution (η^*, x^*, α^*) of the above problem (P). In the next section

(Section 3) a set of necessary optimality conditions is developed in the form of a boundary value problem and an iterative scheme is proposed to obtain a local minimum of J. We illustrate the method developed in Section 3 by a numerical example in Section 4. Finally some extensions of the results in the previous sections and discussions are given in Section 5.

2. EXISTENCE OF SOLUTION FOR PROBLEM (P)

In this section we present some sufficient conditions for the existence of a solution of the optimization problem (P). As it will be seen from the following example the problem (P) as such need not always have a solution.

Let the system equations be given by,

$$\dot{x}(t) = \frac{1}{1+\alpha}, \qquad t \geq 0 \tag{2.1a}$$

$$x(0) = 1 \tag{2.1b}$$

and the functional J be,

$$J = \int_0^T (1 - x(t))^2 \, dt$$

Then we obtain,

$$J = \frac{T^3}{3} \frac{1}{(1+\alpha)^2}$$

Since $\lim_{\alpha \to \pm\infty} J = 0$, there is no finite value of α for which J attains its infimum.

This simple problem illustrates the need to ensure the existence of a solution of the problem (P). It is clear that we need to impose certain restrictions, additional to the mere existence of a solution of the initial value problem (2.1), on the function f and the functional J. In this connection we present below two theorems which assures the existence of a solution of the problem (P).

Theorem-1 :

Let the function f and the functional J in problem (P) satisfy the following conditions :

(i) The function $f(t,x,y,\alpha)$ is continuous with respect to all its arguments, Lipschitzian in (x,y,α); i.e.,

$\exists\ L > 0 \ni$ for every (t,x_1,y_1,α_1) and (t,x_2,y_2,α_2) in the domain of f,

$$|f(t,x_1,y_1,\alpha_1) - f(t,x_2,y_2,\alpha_2)| \leq L(|x_1-x_2| + |y_1-y_2| + |\alpha_1-\alpha_2|)$$

($|\cdot|$ denotes the Euclidian norm)

and is such that there is a solution of the initial value problem (1.1) in $[0,T]$ for all (μ,α).

(ii) The functional J is continuous in all its arguments and satisfies the following conditions :

a) J is bounded from below

b) $J \to \infty$ as $|\alpha| \to \infty$

c) $J \to \infty$ as $|\mu| \to \infty$

Then $\exists\, (\mu^*, x^*, \alpha^*)$ determining a solution of the optimization problem (P).

Proof :

Let,

$$\beta := \inf.\ J\ (\mu, x, \alpha)$$

where the infimum is taken over all (μ, x, α) satisfying equation (1.1). Then $\exists$ a sequence $\{\mu^i, x^i, \alpha^i\}_{i \in N}$ in $R^{nxl} \times C_n[0,T] \times R^m$ satisfying the system equations (1.1) $\ni J^i (:= J(\mu^i, x^i, \alpha^i))$ is monotonic decreasing and $\lim\limits_{i\to\infty} J^i = \beta$. Since J is bounded from below, we have,

$$-\infty < \beta \leqslant J^i \leqslant J^o \ \forall\ i \in N .$$

From conditions (iib) , (iic) and the above inequality we get that $\{\mu^i\}_{i \in N}$ and $\{\alpha_i\}_{i \in N}$ are bounded.

By equation (1.1) and using condition (i) for $o \leqslant t \leqslant T$ we have,

$$|x^i(t)| \leqslant |\mu^i\ e(o)| + \int_o^t |f(s, x^i(s),\ x^i(s-1), \alpha^i)$$

$$- f(s,o,o,o)|ds \quad + \left| \int_o^t f(s,o,o,o)\ ds \right|$$

$$\leqslant |\mu^i\ e(o)| + T.\ \underset{t \in [0,T]}{\mathrm{Max}} |f(t,o,o,o)|$$

$$+ L \int_o^t (|x^i(s)| + |x^i(s-1)| + |\alpha^i|)\ ds$$

$$\leqslant |\mu^i\ e(o)| + T.\underset{t \in [0,T]}{\mathrm{Max.}} |f(t,o,o,o)| + L.T.|\alpha^i|$$

$$+ L \int_o^1 |x^i(s-1)|\ ds + L \int_o^t |x^i(s)|\ ds + L \int_o^t |x^i(s)|\ ds$$

$$\leq |\mu^i e(o)| + T. \underset{t\in[0,T]}{\text{Max.}} |f(t,o,o,o)| + L.T. |\alpha^i| + L.|\mu^i|\cdot||e||$$

$$+ 2 L \int_o^t |x^i(s)| \, ds$$

$$\leq M(\mu^i,\alpha^i) + 2L \int_o^t |x^i(s)| \, ds$$

Obviously, M is a continuous function of α^is and μ^is. Since α^is and μ^is are bounded;

$$M(\mu^i,\alpha^i) \leq \overline{M} < \infty .$$

Thus,

$$\overline{M} \geq |\mu^i e(o)| + T.\underset{t\in[0,T]}{\text{Max.}} |f(t,o,o,o)| + L.T. |\alpha^i|$$

$$+ L |\mu^i| \cdot ||e||$$

By Gronwall's Lemma,

$$|x^i(t)| \leq \overline{M} e^{2 LT} \quad \forall \; i\in N \wedge t\in[0,T]$$

Further,

$$\left|\frac{dx^i(t)}{dt}\right| = | f(t,x^i(t), x^i(t-1), \alpha^i)| \leq \tilde{M} \;\forall\; t\in [0,T]$$

since (μ^i,x^i,α^i) are bounded and f is continuous with respect to all its arguments.

Thus $\{\mu^i,x^i,\alpha^i\}_{i\in N}$ is uniformly bounded and $\{x^i\}_{i\in N}$ is equicontinuous. Ascoli's lemma ensures that a subsequence $\{x^{i_k}\}_{i_k\in N}$ of $\{x^i\}_{i\in N}$ converges uniformly in $[0,T]$. Taking the corresponding subsequence $\{\mu^{i_k}\}_{i_k\in N}$ and $\{\alpha^{i_k}\}_{i_k\in N}$ of $\{\mu^i\}_{i\in N}$ and $\{\alpha^i\}_{i\in N}$ respectively, we have that the subsequence

$\{\mu^{i_k}, x^{i_k}, \alpha^{i_k}\}_{i_k \in N}$ converges. Denote the limit of the subsequence by (μ^*, x^*, α^*). Now taking limits on both sides of the equations ;

$$x^{i_k}(t) = \mu^{i_k} e(o) + \int_0^t f(s, x^{i_k}(s), x^{i_k}(s-1), \alpha^{i_k})\, ds, \quad t \in [0,T]$$

$$x^{i_k}(t) = \mu^{i_k} e(t), \quad t \in [-1,0]$$

We obtain,

$$x^*(t) = \mu^* e(o) + \int_0^t f(s, x^*(s), x^*(s-1), \alpha^*)\, ds, \quad t \in [0,T]$$

$$x^*(t) = \mu^* e(t), \quad t \in [-1,0]$$

Hence, (μ^*, x^*, α^*) satisfies the system equations (1.1) and similarly, $J(\mu^*, x^*, \alpha^*) = \beta$ is ensured by continuity of J with respect to its arguments.

This proves that (μ^*, x^*, α^*) determines an optimal solution of the problem (P). ✓

In the next theorem we consider the minimization of the functional J of the problem (P) when the initial function η is a given (fixed) element of $C_n[-1,0]$. This turns out to be a special case of the problem (P) and we need to ensure the existence of the optimal solution (x^*, α^*) that minimizes the functional J for a fixed $\eta = \bar{\eta} \in C_n[-1,0]$ and satisfies equations (1.1).

<u>Theorem 2</u> :

Let the function f and the functional J of problem (P) satisfy the following conditions :

(i) The function $f(t,x,y,\alpha)$ is continuous with respect to all its arguments and Lipschitzian in (x,y,α); i.e., $\exists L > o \ni$ for every (t,x_1,y_1,α_1) and (t,x_2,y_2,α_2) in the domain of f,

$$|f(t,x_1,y_1,\alpha_1) - f(t,x_2,y_2,\alpha_2)| \leq L(|x_1-x_2| + |y_1-y_2| + |\alpha_1-\alpha_2|)$$

and is such that there is a solution of the initial value problem (1.1) in $[0,T]$ with $\eta = \bar{\eta}$ and all α.

(ii) The functional J is continuous in all its arguments, bounded from below and $|\alpha| \to \infty \implies J \to \infty$

Then $\exists (x^*,\alpha^*)$, $x^* \in C_n[0,T]$ and $\alpha^* \in R^n$ such that J attains a minimum among all (x,α) satisfying equations (1.1) for a fixed $\eta(= \bar{\eta} \in C_n[-1,0])$.

<u>Proof</u> :

Let,

$$\beta := \inf.\ J(x,\alpha)$$

where the infimum is taken over all (x,α) satisfying equations (1.1) with $\eta = \bar{\eta}$. Then $\exists$ a sequence $\{x^i,\alpha^i\}_{i \in N}$ in $C_n[0,T] \times R^m$ satisfying the system equations (1.1) with $\eta = \bar{\eta} \ni J^i (:= J(x^i,\alpha^i))$ is monotonic decreasing and $\lim\limits_{i\to\infty} J^i = \beta$. Then by similar arguments as that of Theorem-1, one obtains a convergent subsequence $\{x^{i_k}, \alpha^{i_k}\}_{i_k \in N}$, whose limit (x^*,α^*) satisfies the integral equation,

$$x^*(t) = \bar{\eta}(o) + \int_o^t f(s,x^*(s), x^*(s-1),\alpha^*)\, ds,\ t \in [0,T]$$

$$x^*(t) = \bar{\eta}(t),\ t \in [-1,0]$$

Hence, (x^*, α^*) satisfies the system equations (1.1) with initial function $\eta = \bar{\eta}$ and $J(x^*,\alpha^*) = \beta$.

3. NECESSARY CONDITIONS

In this section a set of necessary conditions is developed in the form of a boundary value problem, to be satisfied by a solution (η^*,x^*,α^*) of the problem (P). Then a method is proposed to determine a local minimal solution of the problem using these conditions.

We shall impose the following restrictions on the function f and the functional J for the derivation of the necessary conditions :

(i) The function f is continuous in t, twice continuously differentiable with respect to (x,y,α) and is such that there is a solution of the initial value problem (1.1) in $[0,T]$ for all (μ,α).

(ii) The functional J has an Integral representation,

$$J = \int_0^T g(t,x(t),\ x(t-1),\ \alpha)\ dt \quad \text{and}$$

$g(t,x,y,\alpha)$ is continuous in t and twice continuously differentiable with respect to (x,y,α).

Now let,

$$z(t) := \int_0^t g(s,x(s),\ x(s-1),\ \alpha)\ ds \tag{3.1}$$

Making use of the continuous differentiability of the solution [15] $x \in C_n[0,T]$ of the system equations (1.1) with respect to the parameters μ and α it is seen that $z(T,\alpha,\mu)$ is continuously differentiable with respect to (μ,α). So the necessary conditions for optimality are,

$$\frac{\partial z}{\partial \mu_{ij}}(T,\mu,\alpha) = 0, \quad i = 1,\ldots, n \;;\; j = 1, \ldots, m \tag{3.2a}$$

$$\frac{\partial z}{\partial \alpha_i}(T,\mu,\alpha) = 0, \quad i = 1, \ldots, m \tag{3.2b}$$

In order to make use of the above necessary conditions to determine the optimal parameters, we consider the following system of differential equations :

$$\frac{dx}{dt}(t) = f(t,x(t),x(t-1),\alpha) \tag{3.3a}$$

$$\frac{d}{dt}\left(\frac{\partial x(t)}{\partial \mu_{ij}}\right) = \left[\frac{\partial f}{\partial x}\right] \frac{\partial x(t)}{\partial \mu_{ij}} + \left[\frac{\partial f}{\partial y}\right] \frac{\partial x(t-1)}{\partial \mu_{ij}} \tag{3.3b}$$

$$\frac{d}{dt}\left(\frac{\partial x(t)}{\partial \alpha_k}\right) = \left[\frac{\partial f}{\partial x}\right] \frac{\partial x(t)}{\partial \alpha_k} + \left[\frac{\partial f}{\partial y}\right] \frac{\partial x(t-1)}{\partial \alpha_k} + \frac{\partial f}{\partial \alpha_k} \tag{3.3c}$$

$$\frac{d}{dt}\left(\frac{\partial z(t)}{\partial \mu_{ij}}\right) = \left[\frac{\partial g}{\partial x}\right]' \frac{\partial x(t)}{\partial \mu_{ij}} + \left[\frac{\partial g}{\partial y}\right]' \frac{\partial x(t-1)}{\partial \mu_{ij}} \tag{3.3d}$$

$$\frac{d}{dt}\left(\frac{\partial z(t)}{\partial \alpha_k}\right) = \left[\frac{\partial g}{\partial x}\right]' \frac{\partial x(t)}{\partial \alpha_k} + \left[\frac{\partial g}{\partial y}\right]' \frac{\partial x(t-1)}{\partial \alpha_k} + \frac{\partial g}{\partial \alpha_k} \tag{3.3e}$$

for $t \in [0,T]$ and $i = 1, \ldots, n \;;\; j = 1, \ldots, l; \; k = 1, \ldots, m.$

If (μ^*, x^*, α^*) determines a solution of the optimization problem (P) and the function f and the functional J satisfy conditions (i) and (ii) respectively. Then (μ^*, x^*, α^*) must satisfy the system of differential equations (3.3) together with the following boundary conditions :

$$\left.\begin{aligned} & x(t) = \mu . e(t) \\ & \frac{\partial x(t)}{\partial \mu_{ij}} = e_j(t) \\ & \frac{\partial x(t)}{\partial \alpha_k} = 0 \end{aligned}\right] \quad \begin{aligned} & -1 \leqslant t \leqslant 0 \\ & i = 1, \ldots, n \\ & j = 1, \ldots, l \\ & k = 1, \ldots, m \end{aligned} \tag{3.4a}$$

$$\left.\begin{aligned} & \frac{\partial z(0,\mu,\alpha)}{\partial \mu_{ij}} = 0 \\ & \frac{\partial z(0,\mu,\alpha)}{\partial \alpha_k} = 0 \end{aligned}\right] \quad \begin{aligned} & i = 1, \ldots, n \\ & j = 1, \ldots, l \\ & k = 1, \ldots, m \end{aligned} \tag{3.4b}$$

$$\left.\begin{aligned} & \frac{\partial z(T,\mu,\alpha)}{\partial \mu_{ij}} = 0 \\ & \frac{\partial z(T,\mu,\alpha)}{\partial \alpha_k} = 0 \end{aligned}\right] \quad \begin{aligned} & i = 1, \ldots, n \\ & j = 1, \ldots, l \\ & k = 1, \ldots, m \end{aligned} \tag{3.4c}$$

The initial conditions (3.4a) are obtained from (1.1b) and (3.4b) are obtained from the integral equation (3.1). The final conditions (3.4c) are the necessary optimality criteria (3.2).

We justify the above assertion as follows :

Let,

$$x(t,\bar{\mu},\bar{\alpha}) = \bar{\mu}.e(o) + \int_o^t f(s,x(s),\ x(s-1),\ \bar{\alpha})\ ds,\ t\in[0,T] \tag{3.5a}$$

$$x(t,\bar{\mu},\bar{\alpha}) = \bar{\mu}\ \ e(t),\ t\in[-1,0] \tag{3.5b}$$

and

$$\mathbf{z}(t,\bar{\mu},\bar{\alpha}) = \int_o^t g(s,x(s),\ x(s-1),\bar{\alpha})\ ds \tag{3.5c}$$

for some $(\bar{\mu},\bar{\alpha}) \in R^{nxl} \times R^m$.

(3.5a) is a solution of (1.1a), with initial function (3.5b), for the parameter values $(\bar{\mu},\bar{\alpha})$. (3.5c) is the equation (3.1) with parameter values $(\bar{\mu},\bar{\alpha})$ and for $t = T$; it represents the functional J for parameters $(\bar{\mu},\bar{\alpha})$. From the differentiability assumptions on f and g and noting that the solutions of (1.1) and (3.1) are continuously differentiable with respect to the parameters (μ,α), by taking partial derivatives of (3.5) with respect to (μ,α) at $(\bar{\mu},\bar{\alpha})$ first and then the total derivative with respect to t, one obtains the system of equations (3.3) along with the initial conditions (3.4a) and (3.4b). Now if $(\bar{\mu},\bar{\alpha})$ is a solution of the problem (P), in addition $z(T,\bar{\mu},\bar{\alpha})$ must satisfy the necessary conditions (3.2) and hence the final conditions (3.4c). Thus the assertion is justified.

As can be seen in system (3.3), (3.4) there are $(n+n^2.l) + n.m + n.l + m)$ equations and $(n+n^2.l+n.m+n.l+m+nl+m)$ boundary conditions. The $(n.l+m)$ excess conditions will help in the determination of the $(n.l+m)$ parameters (μ,α).

It is to be noted that the system of equations (3.3) with initial conditions (3.4a) and (3.4b) posses a unique solution for fixed (μ,α) by the twice continuous differentiability conditions on f and g with respect to (x,y,α).

The above boundary value problem can be solved by augmenting the system equations by the $(n.l + m)$ additional equations,

$$\frac{d\ \mu_{ij}}{dt} = 0, \quad i = 1, \ldots, n\ ;\ j = 1, \ldots, l$$

$$\frac{d\ \alpha_k}{dt} = 0, \quad k = 1, \ldots, m.$$

and then using any available technique for solution of boundary value problems for ordinary differential equations.

Below we propose an iterative scheme to obtain a local minimum of J using the system of equations (3.3) along with the boundary conditions (3.4).

The solution of (3.3) with initial conditions (3.4a) and (3.4b) at $t = T$ determines the gradient of the functional $J(: = J(\mu,\alpha))$ for a given set of parameters (μ,α). Hence any minimization technique using the gradients can be applied to obtain a local minimal solution (x^*,α^*). Here we use a negative gradient method.

To start with choose (μ^o,α^o) such that at least one component of the vector $\frac{\partial z}{\partial \mu}(T,\mu^o,\alpha^o)$ or $\frac{\partial z}{\partial \alpha}(T,\mu^o,\alpha^o)$ is non-zero.

Choose an acceleration factor $C < 0$. Let (μ^i, α^i) be already determined. Set $\mu^{i+1} = \mu^i + C \frac{\partial z}{\partial \mu} (T, \mu^i, \alpha^i)$

$$\alpha^{i+1} = \alpha^i + C \frac{\partial z}{\partial \alpha} (T, \mu^i, \alpha^i)$$

The iteration is stopped if $J(\mu^{i+1}, \alpha^{i+1}) = J(\mu^i, \alpha^i)$ and (μ^i, α^i) determines a local minimum of J.

The acceleration factor C may be varied in any intermediate step depending on the rate of convergence.

4. NUMERICAL EXAMPLE

As an illustration of the above method we consider the following model reduction problem.

For the chemical process discussed in Chapter 4, our interest is to obtain a reduced model of a specified structure and determine corresponding initial function so that the model best approximates the deviation in the weight composition of reactant A from the nominal value for some known system disturbance.

Below we give a mathematical formulation of the problem :

Consider the uncontrolled process; with $\underline{u} = 0$. The system representation is ;

$$\dot{x}(t) = \begin{bmatrix} -4.93 & -1.01 & 0.0 & 0.0 \\ -3.20 & -5.30 & -12.8 & 0.0 \\ 6.40 & 0.347 & -32.5 & -1.04 \\ 0.0 & 0.833 & 11.0 & -3.96 \end{bmatrix} \bar{x}(t)$$

$$+ \begin{bmatrix} 1.92 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.92 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.87 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.724 \end{bmatrix} \bar{x}(t-1)$$

Let the system disturbance be given as ;

$$\bar{x}_1(t) = 0.1, \quad -1 \leqslant t \leqslant o$$

$$\bar{x}_i(t) = 0.0, \quad -1 \leqslant t \leqslant o \; ; \quad i = 2,3,4$$

We consider a first order reduced model for the above system given by ;

$$\dot{x}(t) = \alpha_1 \; x(t) + \alpha_2 \; x(t-1), \quad t \geqslant o$$

$$x(t) = \mu_1 + \mu_2 t, \quad -1 \leqslant t \leqslant o$$

The functional to be minimized can be justifiably chosen as;

$$J = \tfrac{1}{2} \int_0^T (\bar{x}_1(t) - x(t))^2 \, dt;$$ T is chosen sufficiently large.

The system of associated equations is obtained as follows :

$$\frac{dx(t)}{dt} = \alpha_1 x(t) + \alpha_2 x(t-1)$$

$$\frac{d}{dt}\left(\frac{\partial x(t)}{\partial \mu_1}\right) = \alpha_1 \frac{\partial x(t)}{\partial \mu_1} + \alpha_2 \frac{\partial x(t-1)}{\partial \mu_1}$$

$$\frac{d}{dt}\left(\frac{\partial x(t)}{\partial \mu_2}\right) = \alpha_1 \frac{\partial x(t)}{\partial \mu_2} + \alpha_2 \frac{\partial x(t-1)}{\partial \mu_2}$$

$$\frac{d}{dt}\left(\frac{\partial x(t)}{\partial \alpha_1}\right) = \alpha_1 \frac{\partial x(t)}{\partial \alpha_1} + \alpha_2 \frac{\partial x(t-1)}{\partial \alpha_1} + x(t)$$

$$\frac{d}{dt}\left(\frac{\partial x(t)}{\partial \alpha_2}\right) = \alpha_1 \frac{\partial x(t)}{\partial \alpha_2} + \alpha_2 \frac{\partial x(t-1)}{\partial \alpha_2} + x(t-1)$$

$$\frac{d}{dt}\left(\frac{\partial z(t)}{\partial \mu_1}\right) = (x(t) - \bar{x}_1(t)) \frac{\partial x(t)}{\partial \mu_1}$$

$$\frac{d}{dt}\left(\frac{\partial z(t)}{\partial \mu_2}\right) = (x(t) - \bar{x}_1(t)) \frac{\partial x(t)}{\partial \mu_2}$$

$$\frac{d}{dt}\left(\frac{\partial z(t)}{\partial \alpha_1}\right) = (x(t) - \bar{x}_1(t)) \frac{\partial x(t)}{\partial \alpha_1}$$

$$\frac{d}{dt}\left(\frac{\partial z(t)}{\partial \alpha_2}\right) = (x(t) - \bar{x}_1(t)) \frac{\partial x(t)}{\partial \alpha_2}$$

The set of initial conditions are,

$$\left.\begin{array}{l} x(t) = \mu_1 + \mu_2 t \\[1em] \dfrac{\partial x(t)}{\partial \mu_1} = 1 \\[1em] \dfrac{\partial x(t)}{\partial \mu_2} = t \\[1em] \dfrac{\partial x(t)}{\partial \alpha_1} = 0 \\[1em] \dfrac{\partial x(t)}{\partial \alpha_2} = 0 \end{array}\right] -1 \leqslant t \leqslant o$$

$$\frac{\partial z(o)}{\partial \mu_1} = \frac{\partial z(o)}{\partial \mu_2} = \frac{\partial z(o)}{\partial \alpha_1} = \frac{\partial z(o)}{\partial \alpha_2} = 0$$

We choose T to be 9 units (= 90 mts).

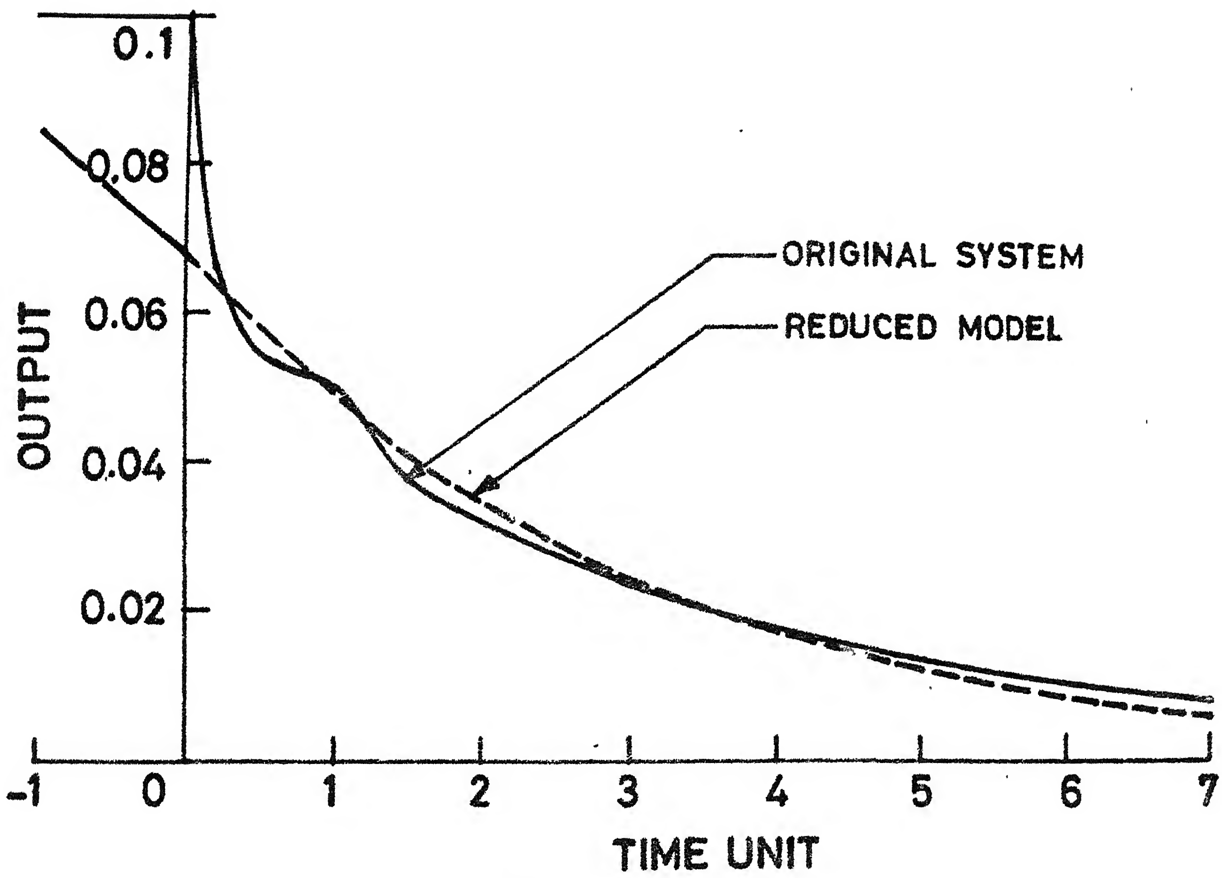

FIG. V.1 DEVIATION IN THE WEIGHT COMPOSITION OF REACTANT - A

Using the iteration scheme presented in the previous section with constant acceleration factor C = -1 and initial guess,

$$\mu_1^o = 0.1, \quad \mu_2^o = 0.0, \quad \alpha_1^o = -0.1, \quad \alpha_2^o = 0.0$$

We obtain,

$$\bar{\mu}_1 = 0.06816877 , \qquad \bar{\mu}_2 = -0.01644707$$

$$\bar{\alpha}_1 = -0.2113325 , \qquad \bar{\alpha}_2 = -0.09254916$$

for a gradient,

$$\frac{\partial z(T)}{\partial \mu_1} = 0.0000248, \qquad \frac{\partial z(T)}{\partial \mu_2} = -0.0001132$$

$$\frac{\partial z(T)}{\partial \alpha_1} = -0.0000547, \qquad \frac{\partial z(T)}{\partial \alpha_2} = -0.0000906$$

and the functional value is obtained as ; J = 0.00004000576.

The system response is shown in figure (V.1).

5. CONCLUSION

In this chapter we have considered the problem of initial function and parameter determination for a system of DDE with single constant delay, minimizing a given functional of system variables. The results can be extended directly to multiple time varying delays when the delays vary continuously with time. The method can be generalized to more complex systems as well. In the derivation of the results the function f is assumed to be

continuous in t. With suitable modifications in the derivations the class of functions f can be extended to incorporate functions satisfying Caratheodory assumptions in t.

The necessary conditions are derived for $\eta \in X \subset C_n[-1,0]$; X a finite dimensional subspace of $C_n[-1,0]$. A near-optimal solution in $C_n[-1,0]$ is obtained if the dimension of X is chosen sufficiently, large. The increase in dimension of X increases the number of parameters μ_{ij} which on the other hand increases the order of the system of associated equations to be solved in each step of the iteration scheme for the determination of the gradient of J. Thus the dimension of the subspace X cannot be increased arbitrarily as the computation time required in that case will be enormous. The choice of the set of vectors $\{e_i : i = 1,2, \ldots, l\}$ in such problems plays an important role, as for the same dimension of the subspace X, one obtains different minimum values of J for different sets of vectors $\{e_i : i = 1,2, \ldots, l\}$. Thus the set of vectors $\{e_i : 1,2,\ldots,l\}$ should be judiciously chosen depending on the nature of the system.

REFERENCES

1. McSchane, 'On multipliers for Lagrange Problems', American Jl. of Math. 61(1939), pp. 809-819.

2. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mishchenko, E.F., 'The mathematical theory of optimal processes', John Wiley and Sons, Inc., New York, 1962.

3. Dubovitskii, A. Ya., Milyutin, A.A., 'Extremum problems in the presence of restrictions', Zh. vychisl. Mat. Mat. Fiz. 5, 3(1965), pp. 395-453.

4. Gamkrelidze, R.V., 'On some extremal problems in the theory of differential equations with application to the theory of optimal control', SIAM Jl. of Cont. Ser A, vol. 3, No.1 (1965), pp 106-128.

5. Neustadt, L.W., 'An abstract variational theory with application to a large class of optimization problems I - General Theory', SIAM Jl. of Cont. Vol. 4, No.3 (1966), pp. 505-527.

6. Neustadt, L.W., 'A general theory of extremals', Jl. of Comp. and Sys. Sc.: 3 (1969), pp. 57-92.

7. Bazaraa, M.S., Goode, J.J., 'Necessary optimality criteria in mathematical programming in normed linear spaces', Jl.of Opt. Th. and Appl., vol. 11, No.3 (1973).

8. Das, P.C., 'Constrained optimization problems in Bannach spaces', Jl. of Opt. Th. and Appl., vol. 17, No.3/4 (1975), pp. 279-291.

9. Lasiecka, I., 'Generalization of Dubovitzky - Milutyn optimality conditions', Jl. of Opt. Th. and Appl., vol. 24, No.30(1978), pp. 421-436.

10. Dubovitskii, A. Ya., Milyutin, A.A., 'Second variations in extremal problems with constraints', Soviet. Math. Doklady 160(1965), 1, pp 18-21.

11. Hoffman, K.H., Kornstaedt, H.J., 'High order necessary conditions in abstract mathematical programme', Free University of Berlin, Berlin (1977).

12. Levitin, E.S., Milyutin, A.A., Osmolovskii, N.P., 'On necessary and sufficient conditions for a local minimum in a problem with constraints', Soviet. Mat. Dokl., vol. 14 (1973), No.3.

13. Kantorovich, L.V., Mathematical methods in organization and planning of production (in Russian), Leningrad, Leningrad State University, (1939).

14. Coddington, E.A., Levinson, N., 'Theory of Ordinary Differential Equations', Tata McGraw-Hill Publishing Co. Ltd., Bombay - New Delhi (1955).

15. Elsgolts, L.E., Norkin, S.B., 'Introduction to the theory and application of differential equations with deviating arguments', 105, Mathematics in Science and Engineering Academic Press (1973).

16. Miller, R.K., 'Nonlinear Volterra Integral Equations', W.A. Benjamin, Inc., California (1971).

17. Falb, P.L., Jan. L. de Jong, 'Some successive approximation methods in control and oscillation theory', 59, Mathematics in Science and Engineering, Academic Press, New York (1969).

18. Bellman, R.E., Kalaba, R.E., 'Quasilinearization and Nonlinear Boundary - Value Problems', The RAND Corporation, 1700 Main St., Santa Monica, California (1965).

19. Hoffman, K.H., Klostermair, A., 'Approximation with solutions of Differential Equations' (in German), 556 Lecture Notes in Mathematics, Springer Verlag.

20. Bard, Y., 'Nonlinear Parameter Estimation', Academic Press, New York (1974).

21. Junkins, J.L., 'An introduction to optimal estimation of dynamical systems', Sijthoff and Noordhoff Int. Publishers B.V., Alphen aan den Rijn, The Netherlands (1978).

22. Davison, E.J., 'A method for simplifying linear dynamic systems', IEEE Tr. Aut. Cont., Vol. 11 (1966), pp. 93-101.

23. Anderson, J.H., 'Geometrical approach to reduction of dynamical systems', Proc. IEE Vol. 114 (1967), pp. 1014-1018.

24. Wilson, D.A., 'Optimal solution of model reduction problem', Proc. IEE, vol. 117(1970), pp. 1161-1165.

25. Iwai, Z., Kubo, Y., 'Determination of eigenvalues in Marshal's model reduction', Int. Jl. of Cont., vol. 30, No.5 (1979), pp. 823-836.

26. Mishra, R.N., Wilson, D.A., 'A new algorithm for optimal reduction of multivariable systems', Int. Jl. of Cont., vol. 31, No.3 (1980), pp. 443-446.

27. Friedman , A., 'Optimal control for hereditary processes', Archive for Rational Mechanics and Analysis, vol. 15 (1964).

28. Vinokurov, V.R., Optimal control of processes described by Integral Equations - I. SIAM Control, vol. 7, No.2 (1969).

29. Huang, S.C., 'Optimal control problems with a system of Integral Equations and restricted phase coordinates', SIAM Control, vol. 10, No.1 (1972), pp. 14-36.

30. Das, P.C., 'The maximum principle of Pontryagin for hereditary processes with constant time-lag and parameters' (In Russian), Trudi Seminar po differentialnikh uravneni, Moscow, Vol. IV (1967).

31. Eller, D.H., Agarwal, J.K., Banks, H.T., 'Optimal control of linear time-delay systems', IEEE Tr. Aut. Cont., vol.AC-14 (1969), pp. 678-687.

32. Agarwal, J.K., 'Computation of optimal control for time-delay systems', IEEE Tr. Aut. Cont. vol. AC-15 (1970), pp 683-685.

33. Slater, G.L., Wells, W.R., 'On the reduction of optimal time-delay systems to ordinary ones', IEEE Tr. Aut. Cont. vol. AC-17, (1972), pp. 154-155.

34. Inoue, K., Akashi, H., Ogino, K., Sawaragi, Y., 'Sensitivity approaches to optimization of linear systems with time-delay. Automatica, vol. 7 (1971), pp. 671-679.

35. Jamshidi, M., Malek-Zavarei, M., 'Sub-optimal design of linear control systems with time delay', Proc. IEE, vol. 119(1972), pp. 1743-1746.

36. Malek-Zavarei, M., 'Near optimum design of nonstationary linear systems with state and control delay', Jl. of Opt. Th. and Appl. vol. 30, No.1, (1980).

37. Malek-Zavarei, M., 'Suboptimal control of systems with multiple delays', Jl. of Opt. Th. and Appl., vol.30, No.4 (1980).

38. Gracovetsky, S.A., Vidyasagar, M., 'A simple iterative method for suboptimal control of linear time-delay systems with quadratic cost', Int. Jl. of Cont. vol. 16, No.5(1972).

39. Gracovetsky, S.A., Vidyasagar, M., 'Suboptimal control of neutral systems', Int. Jl. of Cont., vol. 18, No.1 (1973).

40. Vidyasagar, M., Gracovetsky, S.A., 'Iterative quadratic optimization of nonlinear time-lag control systems', Joint Automatic Control Conference, 1973.

41. Ross, D.W., 'Controller design for time-lag systems via a quadratic criterion', IEEE Tr. Aut. Cont. vol. AC-16, No.6 (1971), pp. 664-672.

42. Williams, T.J., Otto, R.E., 'A generalized chemical processing model for the investigation of computer control', AIEE Tr. (Commun. Electron.), vol. 79 (1960), pp. 458-473.

43. Klee, V., 'Separation and support properties of convex sets - a survey', Control Theory and the Calculus of Variations (book) edited by A.V. Balakrishnan, Academic Press (1969).

44. Kharatishvili, G.L., 'The maximum principle in the Theory of Optimal processes with a Delay', Soviet Math.-Doklady, 2, 28-32 (1961).

EE - 1981 - D - PAT - OPT